



universität  
wien

# MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Live-Untertitelung mittels Respeaking:  
eine Workflow-Analyse“

verfasst von / submitted by

Andreas Pechhacker, BA BA

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of  
Master of Arts (MA)

Wien, 2022 / Vienna 2022

Studienkennzahl lt. Studienblatt /  
degree programme code as it appears on  
the student record sheet:

A 070 331 357

Studienrichtung lt. Studienblatt /  
degree programme as it appears on  
the student record sheet:

Masterstudium Translation Deutsch Portugiesisch

Betreut von / Supervisor:

Univ.-Prof. Dr. Franz Pöchhacker



## **Danksagung**

An dieser Stelle möchte ich bei all jenen bedanken, die mir im Laufe des Studiums mit Rat und Tat zur Seite gestanden sind sowie mir Durchhaltevermögen und Mut zugesprochen haben.

Mein besonderer Dank gilt meinem Betreuer, Univ.-Prof. Dr. Franz Pöchhacker, der mich mit seiner Expertise und mit seinen Ratschlägen zu jeder Zeit unterstützte.

Außerdem möchte ich mich bei meinen Kolleg\*innen bedanken, die durch ihre Mithilfe einen wichtigen Beitrag für den Erkenntnisgewinn der vorliegenden Masterarbeit geleistet haben.

Ein großes Dankeschön gilt auch meiner Familie und meinen Freund\*innen, die mich während des Studiums motiviert und mich bei jeglichen Vorhaben – im Inland wie im Ausland – unterstützt haben. Im Besonderen möchte ich hier Marlene und Charlotte hervorheben, die mir als Lektorinnen für die Masterarbeit kritisch zur Seite gestanden sind.



# Inhaltsverzeichnis

Abbildungsverzeichnis .....	viii
Einleitung .....	1
1. Untertitelung .....	3
1.1. Audiovisuelle Translation.....	3
1.2. Untertitel: Definition .....	4
1.3. Geschichtlicher Abriss der (SDH-)Untertitel .....	6
1.4. Barrierefreiheit im Fernsehen .....	7
1.5. Zielgruppe von SDH-Untertiteln und Erwartungshaltung.....	9
1.6. Klassifizierung von Untertiteln .....	10
1.6.1. Sprachliche Parameter .....	10
1.6.2. Vorhandene Vorbereitungszeit.....	12
1.6.3. Technische Faktoren .....	13
1.6.4. Untertitelungsmethode .....	13
1.6.5. Sendeformat .....	14
1.6.6. Erzeugung von Untertiteln .....	14
1.7. Richtlinien, Eigenheiten und Problematiken .....	15
2. Live-Untertitelung.....	17
2.1. Definition und Allgemeines.....	17
2.2. Geschichte und Methoden der Live-Untertitelung .....	19
2.2.1. Eingabe über Tastaturen.....	20
2.2.2. Spracherkennung .....	22
2.2.3. Vergleich der Methoden zur Erstellung von Live-Untertiteln .....	24
2.3. Herausforderungen bei der Live-Untertitelung .....	25
3. Respeaking .....	27
3.1. Definition – Respeaking und Schriftdolmetschen .....	27
3.2. Arbeitsweise .....	28
3.2.1. Respeaking-Prozess bei der Live-Untertitelung.....	28
3.2.2. Arbeitswerkzeuge .....	29

3.2.3.	Workflow .....	30
3.2.4.	Korrekturmethode .....	31
3.3.	Respeaking und Simultandolmetschen – eine Gegenüberstellung .....	32
3.3.1.	Gemeinsamkeiten .....	33
3.3.2.	Unterschiede .....	34
3.3.3.	Fazit der Gegenüberstellung .....	36
3.4.	Ausbildung und Kompetenzen .....	37
3.5.	Effort Model nach Gile (2009) beim Respeaking und problemauslösende Faktoren	39
3.6.	Problemauslösende Faktoren beim Respeaking .....	42
3.6.1.	Politische Rede .....	44
3.6.2.	Nachrichtensendung .....	45
3.6.3.	Unterhaltungsprogramm .....	46
3.6.4.	Politische Gesprächsrunde .....	46
3.6.5.	Zusammenfassung .....	46
3.7.	Strategien beim Respeaking bzw. Schriftdolmetschen .....	47
3.8.	Aktuelle Untersuchungen zum Respeaking in der Live-Untertitelung .....	48
4.	Fragestellungen und Methodik .....	49
4.1.	Zugang zum Feld .....	49
4.2.	Aufstellung live-untertitelter Sendungen .....	49
4.3.	Workflow und Problemfaktoren bei der Live-Untertitelung .....	51
5.	Ergebnisse .....	54
5.1.	Aufstellung live-untertitelter Sendungen .....	54
5.2.	Workflow .....	62
5.2.1.	Vorbereitung .....	63
5.2.2.	Live-Sendung .....	69
5.2.3.	Nachbearbeitung .....	72
5.3.	Problemfaktoren bei der Live-Untertitelung .....	74
5.3.1.	Vorbereitung .....	74
5.3.2.	Live-Teil .....	75
5.3.3.	Nachbearbeitung .....	76

5.3.4. Fazit.....	76
Conclusio.....	78
Bibliografie.....	81
Anhang .....	88
Interviewleitfaden.....	88
Transkription der Interviews .....	89
Interviewperson 1 (IP1) .....	89
Interviewperson 2 (IP2) .....	93
Abstract – Deutsch .....	98
Abstract – Englisch .....	99

# Abbildungsverzeichnis

Abbildung 1: Vergleich der Methoden zur Erstellung von Live-Untertiteln (Lambourne 2007, zit. nach Romero-Fresco 2011:15) .....	24
Abbildung 2: Zeitliche Darstellung des Respeaking-Prozesses bei der Live-Untertitelung (Luyckx et al. 2010:4).....	29
Abbildung 3: Ausschnitt des selbst entworfenen Rasters mit den verschiedenen Kategorien. 51	
Abbildung 4: Prozessmodell beim interlingualen Live-Untertiteln (Pöchhacker/Remael 2019:137) .....	51
Abbildung 5: Minutenaufteilung aller Sendungen in Prozent.....	54
Abbildung 6: Aufteilung der Sendungen in reine live/vorproduzierte und semi-live/gemischte Untertitel in Prozent .....	55
Abbildung 7: Aufteilung aller Sendungen nach Programmsparte in Prozent .....	56
Abbildung 8: Minutenaufteilung nach Programmsparte in Prozent.....	57
Abbildung 9: Aufteilung aller Sendungen nach Sendungsformat in Prozent .....	58
Abbildung 10: Aufteilung der Minuten nach Sendungsformat in Prozent.....	59
Abbildung 11: Aufteilung gemischte Untertitel nach Sendungsformat in Prozent.....	59
Abbildung 12: Aufteilung vorproduzierte bzw. semi-live Untertitel nach Sendungsformat in Prozent.....	60
Abbildung 13: Aufteilung reine live-produzierte Untertitel nach Sendungsformat in Prozent.....	61
Abbildung 14: Screenshot des Programms Open Media und einer bestimmten Sendung.....	64
Abbildung 15: Screenshot des Untertitelungsprogramm FAB mit den geladenen Open Media-Inhalten .....	65
Abbildung 16: Screenshot des Spracherkennungsprogramms beim Hinzufügen und Eintrainieren unbekannter Wörter .....	68
Abbildung 17: Screenshot beim Live-Untertiteln im semi-live Modus.....	70
Abbildung 18: Screenshot beim Live-Untertiteln im Respeaking-Modus.....	72



# Einleitung

In den letzten Jahren hat sich das Feld der Translationswissenschaft aufgrund des Auftretens neuer Berufsbilder sehr stark verändert. Ein Beispiel dafür ist etwa das Respeaking für die Live-Untertitelung. Die Arbeitsweise des Respeakings lässt sich in den Bereich des Schriftdolmetschens einordnen. Dabei handelt es sich um die simultane Verschriftlichung von lautsprachlichen und paraverbalen Äußerungen. Die Respeaking-Methode für die Erstellung von (hauptsächlich) intralingualen Live-Untertiteln scheint auf den ersten Blick kaum Gemeinsamkeiten mit dem Simultandolmetschen zu haben, weil sie keinen (Fremd-)Sprachtransfer darstellt. Auf den zweiten Blick lassen sich aber deutliche Verbindungen zum Dolmetschen erkennen. Die Fertigkeiten von Dolmetscher\*innen und Respeaker\*innen weisen große Überlappungen auf. Trotzdem zählt das Respeaking noch zu einem relativ neuen, unerforschten Feld in der Translationswissenschaft. Es ist kaum bekannt, welche Arten von Sendungen überhaupt live untertitelt werden und wie der Workflow dabei aussieht. Die vorliegende Masterarbeit soll daher einen Einblick in die Arbeitsweise von Respeaker\*innen geben.

Dafür werden zuerst die theoretischen Grundlagen des Respeakings beschrieben. Es werden die Themen audiovisuelle Translation, Untertitelung, Zielgruppe, Barrierefreiheit im Fernsehen (Kapitel 1), Live-Untertitelung (Kapitel 2) und Respeaking (Kapitel 3) näher behandelt. Des Weiteren wird in Kapitel 3 auf einen Vergleich zwischen Simultandolmetschen und Respeaking eingegangen, um herauszuarbeiten, dass es sich beim Respeaking um eine Art intralinguales Simultandolmetschen handelt. Zu guter Letzt werden noch die problemauslösenden Faktoren (*problem triggers*) nach Gile (2009) erläutert und auf das Feld des Schriftdolmetschens ausgeweitet sowie schon bekannte Problemauslöser (*respeaking crisis points*; Szarkowska et al. 2019) beim Respeaking aufgezeigt und aktuelle Untersuchungen zur Live-Untertitelung beleuchtet.

Im anschließenden empirischen Teil der Arbeit wird eine Aufstellung der Sendungen vorgenommen, die live untertitelt werden. Anhand des Dienstplans einer Untertitelungsfirma soll herausgefunden werden, welche Arten von Sendungen die Respeaker\*innen besonders häufig behandeln. Dafür wird ein zweiwöchiger Zeitraum untersucht und ein eigens angelegtes Raster für die Kategorisierung der Sendungen herangezogen. Die Sendungen werden in Programmsparte und Sendungsformat nach Maurer et al. (2021) eingeteilt. Des Weiteren wird beschrieben, wie der Workflow bei der Live-Untertitelung vonstattengeht. Ausgehend von dem Prozessmodell beim interlingualen Live-Untertiteln von Pöchhacker/ Remael (2019) wird der

Workflow beim intralingualen Live-Untertiteln in Vorbereitung, Live-Sendung und Nachbearbeitung unterteilt. Wie der Workflow beim Bearbeiten einer Sendung konkret aussieht, soll anhand eines Beobachtungsrasters zur Selbstreflexion festgehalten werden. Der Verfasser der Masterarbeit ist selbst als Live-Untertitler tätig und bringt eigene Erfahrungen mit ein. Die Ergebnisse werden zusätzlich mit den Erkenntnissen eines Expert\*innen-Interviews mit zwei Respeaker\*innen, die die gleiche Sendung bearbeiten, abgeglichen und ergänzt. In weiterer Folge soll ausgewertet werden, mit welchen Problemfaktoren die Respeaker\*innen beim Live-Untertiteln konfrontiert sind.

# 1. Untertitelung

Die vorliegende Arbeit befasst sich mit Live-Untertiteln für gehörlose Menschen und Menschen mit Hörbeeinträchtigung. Die Untertitelung stellt eine Unterkategorie der audiovisuellen Translation dar, die im Folgenden näher beschrieben wird. Allerdings wird aufgezeigt, dass sie nicht mit dem Übersetzen im herkömmlichen Sinne von einer in eine andere Sprache gleichzusetzen ist. In den darauffolgenden Unterkapiteln wird eine Untertitel-Definition und ein kurzer geschichtlicher Überblick der Untertitel für gehörlose Menschen und Menschen mit Hörbeeinträchtigung gegeben. Danach werden die Bereiche Barrierefreiheit im Fernsehen sowie die Zielgruppe für (intralinguale) Untertitel näher beleuchtet. Des Weiteren wird auf die Klassifizierung sowie Richtlinien, Eigenheiten und Problematiken von Untertiteln eingegangen.

## 1.1. Audiovisuelle Translation

Die audiovisuelle Translation schließt mehrere Aspekte ein, die in audiovisuellen Medien wie Fernsehen und Kino verwendet werden. Hierbei gibt es eine Übertragung einer Ausgangssprache in eine Zielsprache, die in gewisser Weise mit Bild und Ton verbunden sind (vgl. Díaz Cintas/ Remael 2014:12). Jüngst (2010:1) versteht unter audiovisuellem Übersetzen „das Übersetzen von Medienformaten, die einen sichtbaren und einen hörbaren Teil haben“, wobei das ursprünglich vorliegende Material verändert wird. Davon bleiben Teile erhalten, werden ergänzt oder kombiniert. Dabei geht es zum Beispiel um die Übersetzung von Filmen oder Computerspielen, aber auch das Dolmetschen hat visuelle und auditive Anteile. Laut Jüngst differenzieren sich die Formen der audiovisuellen Translation durch den Zieltext. Dieser kann „schriftlich sein oder mündlich, schriftlich für eine mündliche Darbietung, schriftlich für eine fixierte mündliche Darbietung“ (Jüngst 2010:2). Zudem kann der Zieltext in derselben oder in einer anderen Sprache verfasst sein wie der Ausgangstext bzw. kann den Ausgangstext auch ergänzen.

Bis in die 1990er Jahre war die Forschung zur audiovisuellen Translation mit viel Arbeit verbunden, weil es schwierig war, passendes Material von Filmfassungen aus verschiedenen Ländern zu bekommen (vgl. Jüngst 2010:2). Ab den 1990er Jahren gewann die audiovisuelle Translation immer mehr Zuwendung in der Forschung der Translationswissenschaft. Das hatte damit zu tun, weil audiovisuelle Medien seit diesem Zeitpunkt durch den technologischen Wandel einen immer größeren Einzug in der Gesellschaft fanden (vgl. Díaz Cintas/ Remael 2014:8;

Jüngst 2010:1f.). Die audiovisuelle Translation wird mittlerweile aufgrund zahlreicher Publikationen, Konferenzen, Kursen und Forschungsprojekten als eigenständiges Fachgebiet in der Translationswissenschaft anerkannt (vgl. Díaz Cintas et al. 2010:11).

In der Forschung unterscheidet man zwischen drei Arten der audiovisuellen Translation: Synchronisation, Untertitelung und Voice-Over (vgl. Díaz Cintas/ Remael 2014:8). Jüngst fügt noch weitere Formen der audiovisuellen Übersetzung dazu: Hörfilme bzw. Filme mit Audiodeskription für blinde und sehbehinderte Menschen sowie das Filmdolmetschen, das im Kinosaal bei Festivals verwendet wird (vgl. Jüngst 2010:1). Die Untertitelung und Synchronisation gehören zu den am häufigsten verwendeten Mitteln im Bereich der audiovisuellen Translation (vgl. Díaz Cintas/ Remael 2014:12). Mittlerweile lassen sich Untertitel in der heutigen Welt kaum mehr wegdenken (vgl. Díaz Cintas/ Remael 2014:8; Jüngst 2010:1f.). Einen groben Überblick über die Geschichte der Filmuntertitelung und Synchronisation in unterschiedlichen Ländern bietet Nagel (2009). Besonders kleinere Sprachgemeinschaften setzen aus Kostengründen eher auf die Untertitelung als auf Synchronisation (vgl. Nagel 2009:31).

Nach und nach entwickelten sich immer wieder neue, unterschiedliche Formen der audiovisuellen Translation, wodurch hybride Formen entstanden, die den jeweiligen Zielgruppen angepasst sind. Zu den Arten der audiovisuellen Translation zählt auch die intralinguale Untertitelung für gehörlose Menschen und Menschen mit Hörbeeinträchtigung<sup>1</sup> (vgl. Díaz Cintas/ Remael 2014:12f.). Zwar stellt diese Form der audiovisuellen Translation (genauso wie die Audiodeskription) in der Regel keinen Transfer einer (fremdsprachlichen) Ausgangssprache in eine Zielsprache dar, allerdings spricht sich die Mehrheit der Forscher\*innen dafür aus, dass dieser Bereich ebenfalls als Teil der audiovisuellen Translation anzusehen ist. Denn im Grunde genommen geht es hierbei auch um die Übertragung von Informationen für eine Zielgruppe, die sonst keinen Zugang zu diesen hätte (vgl. Díaz Cintas/ Remael 2014:12f.). Eine Untertitel-Definition wird im nächsten Unterkapitel gegeben.

## **1.2. Untertitel: Definition**

Díaz Cintas/ Remael (2014:8) definieren Untertitel als einen geschriebenen Text, der meist am unteren Bildschirmrand eingeblendet ist und es ermöglichen soll, den auditiven Informationen folgen zu können. Dies inkludiert neben dem Originalton der Sprecher\*innen ebenfalls diskursive und paralinguistische Elemente. Im Englischen spricht man in diesem Bereich von „subtitles for the hard of hearing“ oder HoH subs beziehungsweise von SDH (subtitles for the deaf

---

<sup>1</sup> Für die vorliegende Arbeit wird der Ausdruck „gehörlose Menschen und Menschen mit Hörbeeinträchtigung“ verwendet, um jene Zielgruppe einzugrenzen, für die intralinguale Untertitel hauptsächlich gedacht sind.

and hard of hearing)“ (Jüngst 2010:123), wenn es um Untertitel für gehörlose Menschen und Menschen mit Hörbeeinträchtigung geht. In der deutschsprachigen Literatur findet man daher auch häufig die Bezeichnung SDH-Untertitel.

Des Weiteren ist es wichtig, dass die drei Hauptkomponenten (das gesprochene Wort, das Bild und die Untertitel) im Einklang zueinander stehen. Die Zuseher\*innen sollen sich auf alle Aspekte gleichzeitig konzentrieren können. Daher müssen Untertitel so synchron wie möglich zum Bild und zum gesprochenen Wort eingeblendet werden, um so eine mehr oder minder große semantische Deckungsgleichheit mit dem Ausgangstext aufzuweisen. Außerdem sollen die Untertitel in zeitlicher Angemessenheit lesbar sein. Allerdings findet beim Untertitelungsprozess ein Wandel von gesprochenen in geschriebenen Text statt. Dies hat zur Folge, dass nicht alles eins zu eins vom Ausgangstext übernommen werden kann, sondern es teilweise zu Auslassungen kommt. Denn Untertitel haben nur eine begrenzte Anzahl von Zeichen, damit sie immer noch lesefreundlich sind. Im Normalfall sind das zwischen 32 und 41 Zeichen pro Zeile und höchstens zwei Zeilen pro Untertitel (vgl. Díaz Cintas/ Remael 2014:9). Auch Gottlieb (2002:193) schreibt, dass die Untertitelung dazu neigt, den originalen Dialog zu verkürzen. Zudem sind Untertitel von den Normen geschriebener Sprache bestimmt.

Aus diesen Gründen vertreten manche Wissenschaftler\*innen den Standpunkt, dass es sich bei der Untertitelung nicht um Übersetzung, sondern um Adaption handelt. Folglich erklärt sich auch, warum die audiovisuelle Translation für so lange Zeit von der Translationswissenschaft nicht beachtet wurde und es sich daher noch immer um ein recht junges Feld in der Wissenschaft handelt (vgl. Díaz Cintas/ Remael 2014:9). Jedoch sprechen sich Díaz Cintas/ Remael (2014:10f.) hinsichtlich der audiovisuellen Translation für eine flexiblere Ansicht aus. Laut ihnen ist die Adaption zu negativ konnotiert. Sie stützen sich daher auf die drei Arten der Translation von Roman Jakobson (1959), der der Meinung ist, dass (sprachliche und nicht-sprachliche) Zeichen durch andere Zeichen ersetzt werden können, die denselben Sachverhalt beschreiben. Obwohl Díaz Cintas/ Remael (2014:9f.) selbst der Ansicht sind, dass diese Auslegung zwar aufgrund des technologischen Wandels angepasst werden müsste, trifft die Aufteilung aber noch immer die wesentlichen Eckpunkte der Translation: intralingual (bzw. Umformulieren in derselben Sprache), interlingual (Übersetzung von einer Sprache in die andere) und intersemiotisch (verbale in nonverbale Sprache).

Auch Jüngst (2010) verweist auf Jakobson (1971) und dessen Unterteilung der Übersetzung. Sie postuliert, dass alle audiovisuellen Arten der Übersetzung – abgesehen von der interlingualen Übersetzung – auf den ersten Blick wenig mit der Übersetzung im klassischen Sinne zu tun haben. Für sie ist die intralinguale Übersetzung eine Art der Bearbeitung, die in derselben

Sprache stattfindet wie der Ausgangstext. Darüber hinaus sind Untertitel für gehörlose Menschen und Menschen mit Hörbeeinträchtigung eine spezielle Form, weil diese aufgrund des Wechsels des Zeichensystems als intersemiotische Übersetzung angesehen werden können (vgl. Jüngst 2010:3). Hierbei wird Gesprochenes in Geschriebenes umgewandelt und paralinguistische Elemente werden verschriftlicht (vgl. Jüngst 2010:123). Gottlieb (2002:191) bezeichnet Untertitel hingegen als intrasemiotisch, weil sie sich innerhalb der Grenzen der audiovisuellen Medien abspielen.

Intralinguale Untertitel für gehörlose Menschen und Menschen mit Hörbeeinträchtigung sind seltener anzutreffen als interlinguale Untertitel für hörendes Publikum. Allerdings nimmt das Bewusstsein der Öffentlichkeit für SDH-Untertitel immer stärker zu (vgl. Jüngst 2010:2). Ein kurzer geschichtlicher Überblick über die Entstehung von Untertiteln soll das nun beleuchten.

### **1.3. Geschichtlicher Abriss der (SDH-)Untertitel**

Schon im Jahre 1909 wurden Untertitel mithilfe eines Diaprojektors auf Filmleinwände projiziert. Da es zu dieser Zeit nur Stummfilme gab, halfen die Untertitel dabei, den Hauptelementen der Handlung und wichtigen Dialogen besser folgen zu können. Im Grunde genommen hatte diese Art der Untertitel viel mit der SDH-Untertitelung gemein. Allerdings war die Methode mithilfe von Projektoren nur schwer umzusetzen und wurde im Laufe der Zeit von anderen Techniken abgelöst. Trotzdem wird sie heutzutage auch noch bei der Live-Untertitelung im Kino benutzt, wo die Untertitel live mithilfe eines Videobeamers auf die Leinwand projiziert werden. Nach und nach entwickelten sich immer bessere Methoden, die vor allem die Lesefreundlichkeit von Untertiteln verbesserten. Heute zählt die Lasertechnologie wegen der guten Lesbarkeit der Filmuntertitel zur besten Art (vgl. Ivarsson/ Carroll 1998:9-18).

Erste Versuche mit Untertiteln für gehörlose Menschen und Menschen mit Hörbeeinträchtigung gab es ab den späten 1940er Jahren in den USA. Den Durchbruch schafften die SDH-Untertitel im Jahre 1971 bei der ersten nationalen Konferenz zum Fernsehen für gehörlose Menschen und Menschen mit Hörbeeinträchtigung, wo über die Technologie für SDH-Untertitel diskutiert wurde. Die erste Sendung mit Untertiteln lief kurz danach. Dabei handelte es sich um offene Untertitel (eine Klassifizierung erfolgt in Kapitel 1.6.). Erst ab 1980 fand man geschlossene Untertitel, die man mithilfe eines Decoders sichtbar machen konnte. In Europa gab es erst Ende der 70er Jahre bzw. Anfang der 80er Jahre erste Versuche mit Fernsehuntertiteln. Dies hatte auch mit der Einführung des Teletextes zu tun (Romero-Fresco 2011:6f.).

Zwar setzt man in Deutschland vermehrt auf Synchronisation von Filmen und Serien, doch kommt die Verwendung von Untertiteln immer mehr zur Anwendung. Zum einen basiert das auf der Tatsache, dass Untertitel deutlich kostengünstiger und schneller als Synchronisationen produziert werden können. Zum anderen liegt es an sprachpolitischen und publikumsbezogenen Aspekten. Dies hat unter anderem damit zu tun, dass es vermehrt zu gesetzlichen Auflagen kommt, die die Fernsehsender dazu verpflichten, eine bestimmte Quote ihres Programms mit Untertiteln für gehörlose Menschen und Menschen mit Hörbeeinträchtigung bereitzustellen. Außerdem finden Untertitel immer mehr Verwendung beim Erlernen einer Fremdsprache und tragen zum Erhalt von Minderheitensprachen bei. Aus diesen Gründen werden immer mehr Fernsehsendungen untertitelt. Diese reichen von Dokumentarfilmen über Reality-Shows bis hin zu Promi-Magazinen (vgl. Nagel 2009:35f.).

Trotzdem gibt es von Seiten der Gehörlosenbünde Kritik an der Qualität von (Live-)Untertiteln und sie fordern mittels Petitionen und Demonstrationen mehr Untertitel im Fernsehen. Denn es werden nur wenige als wichtig angesehene Sendungen regelmäßig untertitelt. Im April 2009 wurden im deutschen Fernsehen nur zehn Prozent der Sendungen untertitelt (vgl. Jüngst 2010:123). Mittlerweile ist aber davon auszugehen, dass die öffentlich-rechtlichen Rundfunkanstalten ihre Untertitelquoten deutlich gesteigert haben. Denn da auch Menschen mit Sinnesbehinderungen den Rundfunkbeitrag bezahlen müssen, haben die Fernsehanstalten auch dafür zu sorgen, dass diesen Menschen die Teilhabe am gesellschaftlichen Leben ermöglicht wird (vgl. Mälzer/ Wünsche 2020:327f.).

Barrierefreiheit im Fernsehen ist somit ein wichtiger Aspekt für gehörlose Menschen und Menschen mit Hörbeeinträchtigung. Im nächsten Unterkapitel wird dies näher beleuchtet.

#### **1.4. Barrierefreiheit im Fernsehen**

Unter Barrierefreiheit versteht man die Idee des vollen Zugangs zur globalisierten Welt als Menschenrecht. Es geht hier demnach nicht nur um die Überwindung physischer Barrieren, sondern auch um den Zugang zu Informationen (vgl. Remael 2012:95). In vielen Ländern gibt es bereits verpflichtende Vorgaben zur Bereitstellung von (Live-)Untertiteln, welche die Fernsehanstalten umzusetzen haben (vgl. Remael 2012:98).

Im Folgenden werden zwei der bedeutendsten Rechtsakte für Europa vorgestellt, die zu einer Verbesserung der Barrierefreiheit führen sollen: Zum einen ist das die UN-Behindertenrechtskonvention, zum anderen die EU-Richtlinie über audiovisuelle Mediendienste. Durch die Unterzeichnung der UN-Behindertenrechtskonvention (Übereinkommen der Vereinten Nationen über die Rechte von Menschen mit Behinderungen, UN-BRK) wird vermehrt Druck auf

die Mitgliedstaaten ausgeübt. Die Konvention verpflichtet sie, die barrierefreie Rezeption von Informationen auszubauen (vgl. Jekat 2014:89; UN-BRK 2008). Auch die Europäische Union setzt sich für die Verbesserung der Barrierefreiheit in den Medien ein und fordert in Europa Untertitel für möglichst alle Sendungen des öffentlich-rechtlichen Fernsehens (vgl. Jekat 2014:89; Romero-Fresco 2011:10). Um die Teilhabe von Menschen mit Beeinträchtigung am gesellschaftlichen und kulturellen Leben zu sichern, werden Staaten dazu aufgefordert, Inhalte für sie bereit zu stellen. Im Paragraf 22 der EU-Richtlinie zur Koordinierung bestimmter Rechts- und Verwaltungsvorschriften der Mitglieder über die Bereitstellung audiovisueller Mediendienste heißt es etwa:

Das Recht von Menschen mit Beeinträchtigungen und von älteren Menschen auf Teilnahme am gesellschaftlichen und kulturellen Leben der Union und ihre diesbezügliche Integration ist mit der Bereitstellung barrierefreier audiovisueller Mediendienste verbunden. Daher sollten die Mitgliedstaaten ohne unangemessene Verzögerung sicherstellen, dass sich die ihrer Rechtshoheit unterworfenen Mediendienstanbieter aktiv darum bemühen, ihre Inhalte für Menschen mit Behinderungen, insbesondere für Menschen mit Seh- oder Hörstörungen, zugänglich zu machen. (EU 2018)

Die Richtlinie sieht zudem in Paragraf 23 vor, dass Gebärdensprache, Untertitelung für Gehörlose und Schwerhörige, gesprochene Untertitel und Audiobeschreibung angeboten werden sollen. Zwar ließ sich seit dem Inkrafttreten des Bundes-Behindertengleichstellungsgesetzes im Jahr 2002 auch ein Anstieg des Angebots an Untertiteln im Fernsehen feststellen, allerdings ist die Untertitelungsrate im deutschsprachigen Fernsehen noch sehr gering (vgl. Hezel 2009:161f.). Die deutschen Fernsehsender versuchen allerdings, ihr Angebot an Untertiteln auszuweiten, um die Ziele zur Verbesserung der Barrierefreiheit im Fernsehen zu erreichen (vgl. Romero-Fresco 2011:10f.). Dass jedoch Live-Untertitel in der Umsetzung aufwendig zu produzieren sind, zeigt sich auch in der Richtlinie über audiovisuelle Mediendienste. So heißt es in Artikel 22 der Richtlinie über audiovisuelle Mediendienste, dass „praktische und unvermeidbare Einschränkungen, die beispielsweise im Fall von live übertragenen Sendungen oder Veranstaltungen eine vollständige Barrierefreiheit verhindern könnten, zu berücksichtigen sind“ (EU 2018). Dies dürfte besonders mit der Technik zusammenhängen. Remael (2012:99f.) stellt fest, dass im Bereich Live-Untertitelung noch einiges an Verbesserung von Nöten ist und Standards entworfen werden müssen, um im Fernsehen die Barrierefreiheit weiter auszubauen.

Der Deutsche Schwerhörigenbund geht derzeit von knapp 16 Millionen Menschen mit Hörbeeinträchtigung in Deutschland aus (vgl. DSB 2021). Angesichts der demografischen Entwicklung, der immer älter werdenden Bevölkerung und der zunehmenden gesetzlichen Vorgaben kann in Zukunft von einem erhöhten Bedarf an Untertiteln ausgegangen werden. Auch die



Nachfrage nach Live-Untertiteln wird somit wachsen und das Tätigkeitsfeld wird für Translator\*innen vermehrt an Bedeutung gewinnen (vgl. Tiittula 2006:487). Auf die Zielgruppe von SDH-Untertiteln und deren Erwartungshaltung wird im folgenden Unterkapitel eingegangen.

### **1.5. Zielgruppe von SDH-Untertiteln und Erwartungshaltung**

SDH-Untertitel werden für eine recht breite, heterogene Zielgruppe erstellt. Dies können von Geburt an gehörlose, späterschwerhörige und schwerhörige Menschen sein (vgl. Jüngst 2010:124). Laut Neves (2009) werden Menschen als gehörlos bezeichnet, sobald es unmöglich für sie ist, gesprochene Informationen allein durch das Hören zu verstehen. Als schwerhörig werden jene Menschen angesehen, die mit permanenten oder schwankenden Hörverlusten zu kämpfen haben und demnach Schwierigkeiten haben, gewisse Töne zu hören. Sie sind daher nicht komplett gehörlos, sondern haben immer noch ein gewisses Restgehör. Somit sind schwerhörige Menschen zwischen gehörlosen und hörenden Menschen einzuordnen (vgl. Neves 2009:154f.). Leicht- und mittelgradig schwerhörige Menschen können mit technischen Hilfsmitteln wie etwa Hörgeräten hören und sind nicht unbedingt auf Untertitel angewiesen (vgl. Hezel 2009:154f.).

Aufgrund der Heterogenität unterscheiden sich die Bedürfnisse innerhalb der Zielgruppe. Dies muss beachtet werden, damit die Untertitel entsprechend angepasst werden. Denn Menschen, die erst im Laufe ihres Lebens gehörlos oder schwerhörig werden, erlernen normalerweise im Gegensatz zu von Geburt an gehörlosen Menschen keine Gebärdensprache. Da sich die Gebärdensprache grammatikalisch stark von der Lautsprache unterscheidet, sind Untertitel für Menschen mit Hörbeeinträchtigung fremdsprachlich. Allerdings befinden sich unter gehörlosen Menschen viele, die problemlos die Schriftform der Lautsprache lesen und verstehen (vgl. Jüngst 2010:124f.).

Untertitel erfüllen somit unterschiedliche Nutzer\*innenzwecke. Eventuell wollen gehörlose Menschen, die die Deutsche Gebärdensprache als Muttersprache haben, ihre Kenntnisse der deutschen Laut- und Schriftsprache verbessern. Zudem können Untertitel auch von anderen Nutzer\*innen verwendet werden wie von Fremdsprachenlerner\*innen und Legasthener\*innen. Jedoch hebt Hezel (2009) hervor, dass die zuletzt genannten Nutzer\*innen aber keine Adressat\*innen sind, weil sie keine Hörbeeinträchtigung haben und somit nicht auf SDH-Untertitel angewiesen sind. Innerhalb der Zielgruppe werden Untertitel wegen ungleicher Lesekompetenzen unterschiedlich gut verstanden. Die Untertitler\*innen müssen daher abwägen, wie viel (un)bearbeiteter Text in wie viel Zeit gesendet werden kann (vgl. Hezel 2009:154f.).

Aus diesem Grund werden Untertitel von den Rezipient\*innen allerdings auch unterschiedlich wahrgenommen. Gehörlose Zuseher\*innen vergleichen meist die Untertitel mit der Länge der jeweiligen Äußerung der Sprecher\*innen. Falls es zu mangelnder Übereinstimmung kommt, vor allem aufgrund gut absehbarer Laute, erweckt das Unzufriedenheit und Misstrauen bei den Rezipient\*innen (vgl. Prillwitz 2001:90, zit. nach Hezel 2009:160). Demnach postuliert Jüngst (2010:127), dass nicht auf eine Übereinstimmung von sprachlichen Äußerungen und Untertiteln verzichtet werden darf).

Des Weiteren zeigt eine Umfrage des ORF (vgl. Hezel 2009:242f.), dass sich die Mehrheit der Zielgruppe für eine Wort-für-Wort-Untertitelung ausspricht. Neves (2009:158) berichtet allerdings von anderen Ergebnissen in Portugal. Dort nimmt die Zielgruppe diese Art von Untertiteln als zu schwierig, zu lang, zu komplex, zu schnell und zu verschachtelt auf. Es kommt also zu unterschiedlichen länderabhängigen Vorlieben in der Zielgruppe (vgl. Romero-Fresco 2015).

Hezel (2009) sieht in den verschiedenen Präferenzen ein Problem. Untertitel sollen wörtlich, aber auch grammatikalisch korrekt sein. Jedoch kann es vorkommen, dass sich diese Forderungen gegenseitig ausschließen. Außerdem kann die Wort-für-Wort-Übertragung aus Platz- und Zeitgründen nicht immer eingehalten werden (vgl. Hezel 2009:155). Wenn der Text zu lang ist, gibt es zu wenig Zeit, die Untertitel zu lesen (vgl. Jüngst 2010:140). Zwar sollen Vereinfachungen so gut wie möglich vermieden werden (vgl. Jüngst 2010:131), aber Veränderungen im Untertitel sollen nicht gleich als Bevormundung angesehen werden. Sie gelten eher der Textoptimierung, was somit nichts mit der Lesekompetenz der Zielgruppe zu tun hat (vgl. Hezel 2009:248).

## **1.6. Klassifizierung von Untertiteln**

Es ist schwierig, eine fixe Klassifizierung von Untertiteln aufzustellen, weil Untertitel eng mit Technologie verknüpft sind und sich diese sehr schnell weiterentwickelt. Nichtsdestotrotz kategorisieren Díaz Cintas/ Remael (2014:13) Untertitel nach diesen fünf Kriterien: sprachliche Faktoren, vorhandene Vorbereitungszeit, technische Faktoren, Untertitelungsmethode sowie Sendeformat. Auf die Kriterien soll im Folgenden kurz eingegangen werden. Als weiteres Kriterium könnte noch hinzugefügt werden, von wem die Untertitel erstellt werden.

### **1.6.1. Sprachliche Parameter**

Sprachliche Parameter gruppieren Untertitel in intralinguale, interlinguale und bilinguale Untertitel:

- Bei intralingualen Untertiteln wird gesprochener Text verschriftlicht und Ausgangs- sowie Zielsprache sind immer ident. Diese Form der Untertitel wird prinzipiell dazu verwendet, um gehörlosen Menschen und Menschen mit Hörbeeinträchtigung einen besseren Zugang zu audiovisuellen Medien zu verschaffen. Zudem zählt die intralinguale Untertitelung zu jener Form, die sich zurzeit aufgrund bestimmter Gesetzgebungen und Regelungen am meisten in der audiovisuellen Kommunikation entwickelt. Daher ist der Prozentsatz von intralingualen Untertiteln in den letzten Jahren rasant gestiegen (vgl. Díaz Cintas/ Remael 2014:14f.).

Intralinguale Untertitel lassen sich häufiger im Fernsehen als im Kino finden (vgl. Jüngst 2010:124) und werden darüber hinaus für pädagogische Zwecke sowie zum Erlernen von Fremdsprachen eingesetzt. Somit können nicht nur gehörlose Menschen und Menschen mit Hörbeeinträchtigung von intralingualen Untertiteln profitieren, sondern auch Immigrant\*innen oder internationale Studierende. Eine zusätzliche Form von intralingualen Untertiteln ist das Karaoke-Singen sowie das Einblenden von Untertiteln in Filmen oder im Fernsehen, sobald Akzente von Sprecher\*innen für die Zuschauer\*innen schwierig zu verstehen sind, obwohl sie dieselbe Sprache sprechen. Eine weitere Kategorie stellen intralinguale Untertitel in Wartestationen von öffentlichen Verkehrsmitteln dar. Dies ermöglicht, die Botschaft von Werbungen oder Nachrichten ohne Ton zu vermitteln (vgl. Díaz Cintas/ Remael 2014:14-17).

- Bei interlingualen Untertiteln geht es um die Übersetzung einer mündlichen Äußerung, die in einer anderen Sprache in Form von Untertiteln verschriftlicht wird, um fremdsprachliche Sendungen für eine jeweilige Landessprache zur Verfügung zu stellen. Demnach wird wie beim intralingualen Untertiteln auch hier gesprochener Text verschriftlicht. Gottlieb (2002:190f.) spricht von einem diagonalen Übersetzungsvorgang, weil nicht nur die Sprache, sondern auch der Modus verändert wird. Bei interlingualen Untertiteln handelt es sich um eine „verknappte Übersetzung des Filmdialoges“ (Jüngst 2010:25). Des Weiteren ist diese Art der Untertitel für hörendes Publikum gedacht und weist aufgrund dessen nicht dieselben Besonderheiten auf, die für intralinguale SDH-Untertitel charakteristisch sind (vgl. Jüngst 2010:25).
- Bilinguale Untertitel werden vorwiegend in Ländern verwendet, in denen mehr als eine Landessprache gesprochen wird. So gibt es beispielsweise in Belgien französische sowie flämische Untertitel. Die Eigenheit hierbei ist, dass beide Sprachen pro Untertitel vertreten

sind. Meist wird eine Sprache pro Zeile angezeigt, was häufig zu Platzproblemen von Untertiteln führt. Bilinguale Untertitel kommen des Öfteren bei internationalen Filmfestivals zum Einsatz, um ein größeres Publikum zu generieren (vgl. Díaz Cintas/ Remael 2014:18f.).

### 1.6.2. Vorhandene Vorbereitungszeit

Je nach Sendung können Untertitel dementsprechend früh, erst relativ kurzfristig oder gar nicht vorbereitet werden. Hierbei unterscheidet man zwischen vorbereiteten Untertiteln, semi-live Untertiteln und reinen Live-Untertiteln.

- Bei vorbereiteten Untertiteln (auch (pre-)prepared subtitles/Offline-Untertitel)<sup>2</sup> handelt es sich um Untertitel, die zur Gänze vor der Ausstrahlung der schon aufgezeichneten Sendung zeitgerecht erstellt werden können. Die vorbereiteten Untertitel entsprechen meist wörtlich dem Originalton, wobei es auch vorkommen kann, dass sie nur auf das Notwendigste reduziert werden, wie etwa bei (übersetzten) Interviews oder Dokumentationen (vgl. Díaz Cintas/ Remael 2014:19). Diese Untertitel sind schon mit den passenden Timecodes<sup>3</sup> versehen (vgl. EBU 2004:10).
- Semi-live Untertitel können im Vorhinein zwar auch vorbereitet werden, allerdings haben sie noch keine Timecodes, weil die Untertitel erst während der Ausstrahlung einer Sendung live geschickt werden. Das Untertitelungsprogramm versetzt die versendeten Untertitel mit den entsprechenden Timecodes. Semi-live Untertitel werden häufig für Nachrichtensendungen benutzt, sofern Text im Voraus verfügbar ist (vgl. Romero-Fresco 2011:12).
- Live- bzw. Echtzeit-Untertitel (auch Online Untertitel) werden live bzw. simultan zur Ausstrahlung einer Sendung produziert. Es ist eine recht neue Form von Untertiteln und sie wird verwendet, sobald es nicht möglich ist, Untertitel im Vorhinein vorzubereiten. Live-Untertitel kommen zum Beispiel bei politischen Reden, Nachrichteninterviews oder Sportprogrammen zum Einsatz. Meist findet Live-Untertitelung intralingual für gehörlose Menschen und Menschen mit Hörbeeinträchtigung statt. Allerdings gibt es auch immer häufiger inter-

---

<sup>2</sup> Im empirischen Teil der Arbeit wird der Begriff „vorproduziert“ für eine konkretere Abgrenzung zu (vorbereiteten) semi-live Untertiteln verwendet.

<sup>3</sup> Anhand der Timecodes kann man die einzelnen Bilder bzw. Frames bestimmen. Der Code (hh:mm:ss:ff) besteht aus acht Zahlen, die sich aus Stunden (h), Minuten (m), Sekunden (s) und Frame-Nummer (f) zusammensetzen. Die Untertitel können somit bildgenau und synchron zum Originalton und -bild angepasst werden (vgl. Nagel 2009:83f.).

linguale Live-Untertitel. Interlingual werden sie oft in Kombination von Dolmetscher\*innen und Stenograf\*innen erstellt. Eines der größten Probleme bei Live-Untertiteln ist, die Synchronität von Originalton und Untertitel einzuhalten. Denn Untertitel können erst erstellt werden, nachdem der Originaltext gesagt wurde. Dies führt automatisch zu Asynchronität von Untertitel und Dialog der Sprechenden, was wiederum für die Zuseher\*innen verwirrend erscheinen kann (vgl. Díaz Cintas/ Remael 2014:19f.). Es gibt unterschiedliche Methoden, um Live-Untertitel zu erstellen. Diese werden in Kapitel 2.2. näher beleuchtet.

### **1.6.3. Technische Faktoren**

Bei den technischen Faktoren ist zwischen offenen und geschlossenen Untertiteln zu unterscheiden, je nachdem, ob sie selbstständig ein- und ausgeschaltet werden können oder nicht.

- Offene Untertitel werden auf dem Bildschirm aller Zuschauer\*innen gezeigt und werden direkt vom Fernsehsignal ausgestrahlt. Sie sind somit als integrales Element des audiovisuellen Programms für alle Zuseher\*innen sichtbar und nicht wegschaltbar (vgl. Díaz Cintas/ Remael 2014:21 und 249). Die offenen Untertitel lassen sich besonders häufig bei interlingualen Untertiteln finden (vgl. Nagel 2009:35f.).
- Geschlossene Untertitel sind im gesendeten Videosignal verschlüsselt und können erst mithilfe eines Decoders aktiviert und sichtbar gemacht werden. Dies lässt sich mit der Fernbedienung oder via Teletext einstellen, sofern Untertitel für das jeweilige Programm vorgesehen sind. Untertitel lassen sich ebenfalls oft bei DVD-Filmen und Serien (de-)aktivieren. Vor Einführung der DVD waren interlinguale Untertitel immer offen, intralinguale Untertitel hingegen geschlossen (vgl. Díaz Cintas/ Remael 2014:21f.).

### **1.6.4. Untertitelungsmethode**

Wie Untertitel schlussendlich im Kino oder im Fernsehen eingeblendet werden, hat sich im Laufe der Zeit stark verändert. Vor allem Aufmachung und Stabilität der Untertitel wurden verbessert. Die heutzutage am meist verwendeten Untertitel werden mittels Laser (im Kino) bzw. elektronisch (im Fernsehen und DVD) ausgestrahlt (vgl. Díaz Cintas/ Remael 2014:22f.).

### **1.6.5. Sendeformat**

Untertitel werden für das Kino, das Fernsehen, für Videos (bzw. Videokassetten), DVDs und für das Internet produziert. Je nach Sendeformat werden Untertitel angepasst, damit die Zuschauer\*innen ausreichend Zeit haben, um die eingeblendeten Untertitel lesen zu können. Um beispielsweise bei Fernsehuntertiteln 70 bis 74 Zeichen lesen zu können und die Information mit dem gezeigten Bild in Verbindung zu bringen, benötigt man durchschnittlich sechs Sekunden. Bei Kinountertiteln bzw. bei DVDs setzt man im Gegensatz dazu eher auf längere Untertitel, weil diese auf einer großen Leinwand einfacher und schneller zu lesen sind als auf einem Fernsehgerät bzw. Zuseher\*innen das Video bei DVDs zurückspulen können, falls sie nicht genug Zeit hatten, die Untertitel zu lesen (vgl. Díaz Cintas/ Remael 2014:23ff.).

### **1.6.6. Erzeugung von Untertiteln**

Als weitere Kategorie könnte der Auflistung von Díaz Cintas/ Remael (2014) noch die Erzeugung von Untertiteln hinzugefügt werden. Denn einerseits gibt es von Menschen erzeugte Untertitel, wie man sie etwa im Kino oder im Fernsehen sieht. Andererseits findet man auch schon auf Onlineplattformen wie YouTube die Möglichkeit, automatisch von einer Maschine generierte Untertitel anzeigen zu lassen. Diese können intralingual und interlingual sein.

Da die automatische Spracherkennung in den letzten Jahren immer mehr an Bedeutung gewann, kann in Zukunft damit gerechnet werden, dass diese Methode eventuell bald im (Live-)Fernsehen für die Erstellung von Untertiteln eingesetzt wird. Der Text würde dann entweder unkorrigiert gesendet werden oder noch von Untertitler\*innen vor dem Senden korrigiert werden (vgl. Romero-Fresco 2019:100). Laut Remael et al. (2014:143f.) muss es sich erst zeigen, ob diese Untertitelungsmethode auch effizienter als von Menschen produzierte Untertitel sein wird. Sie gehen jedoch bei der interlingualen Live-Untertitelung nicht davon aus, dass automatisch generierte Untertitel die von Menschen produzierten Untertitel ablösen werden. Luyckx et al. (2010:4ff.) stellen zudem klar, dass Spracherkennungssoftwares noch nicht gut genug sind, um mündliche Texte direkt transkribieren zu können. Eine direkte Übernahme von mündlich dargebotenen Reden ohne gewisse Adaption ist für Untertitel nicht geeignet. Denn Satzabbrüche, Korrekturen und Hässitation sind in Untertiteln nicht gewünscht. Auch Käber (2016:301f.) sieht in der direkten automatischen Spracherkennung Nachteile. Zum einen benötigt es wochenlanges Training des persönlichen Sprachprofils, damit die Spracherkennung gut funktioniert und sich an die Sprechweise jedes einzelnen anpasst. Dies wäre bei immer wieder wechselnden Gästen und Hintergrundmusik nicht möglich. Zum anderen ist ein Text – sofern er exakt erkannt werden könnte – ohne Interpunktion kaum lesbar und die Textmenge wäre zu

groß zum Lesen. Es kann davon ausgegangen werden, dass auch die Untertitel-Richtlinien schwer einzuhalten wären. Auf diese wird im nächsten Unterkapitel eingegangen.

## **1.7. Richtlinien, Eigenheiten und Problematiken**

Im Fernsehen gibt es Sendungen in verschiedenen Formaten. Laut Hezel (2009) unterscheiden sich diese in Form, Inhalt, Zielgruppe und Sendeplatz. Sie kategorisiert in Anlehnung an Prillwitz (2001:11) Sendeformate folgendermaßen: „Spielfilme, Serien und Reihen, Kindersendungen, Ratgeber und Dokumentation, Magazine und Reportagen, Naturfilm und Tierdokumentation, Religiöse Sendung und Nachrichten“ (Hezel 2009:152). Das Sendeformat hat einen Einfluss auf die Untertitelungsstrategie und vor allem auf die Entscheidung über die Wort-für-Wort-Übertragung der Untertitel. Dementsprechend müssen die Untertitel der jeweiligen Zielgruppe angepasst werden. Allgemein gültige Richtlinien für alle Sendeformate seien daher schwer festzulegen und sollten recht umfassend formuliert sein, um alle Formate miteinzubeziehen (vgl. Hezel 2009:152ff.).

Richtlinien zur Erstellung von SDH-Untertiteln können sich auch je nach Fernsehanstalt und Untertitelungsfirma unterscheiden (vgl. Jüngst 2010:127). Zwar gibt es Empfehlungen von einer Arbeitsgruppe des Deutschen Gehörlosenbundes, doch diese können in der Praxis oft nicht umgesetzt werden (Hezel 2009:148). Denn SDH-Untertitel weisen einige Eigenheiten wie zum Beispiel parasprachliche und nonverbale Elemente auf, die man so bei Untertiteln für hörende Zuseher\*innen nicht findet (vgl. Neves 2005).

Während Untertitel im Kino einfarbig sind, werden Untertitel in Fernsehen und DVDs oft in verschiedenen Farben eingefärbt, um Sprecher\*innen voneinander unterscheiden zu können. Meist werden die Farben Weiß, Gelb, Grün, Zyan und falls nötig Magenta verwendet. Hierbei sollte angemerkt werden, dass Magenta aufgrund schlechter Lesbarkeit eher ungern verwendet wird. Bei einfarbigen Untertiteln werden die Untertitel so platziert, dass sie den Sprecher\*innen eindeutig zuzuordnen sind. Bei Kommentaren im Off wird explizit im Untertitel verdeutlicht, dass sie aus dem Off stammen (vgl. Jüngst 2010:128ff.).

Neben Dialogtext werden auch parasprachliche Elemente bei Untertiteln verwendet, falls diese wichtig für das Verstehen der Handlung sind. Das können zum Beispiel relevante Geräusche (Klingeln des Telefons, Jubel) oder sprachliche Merkmale von Sprecher\*innen wie deren Intonation (Ironie) sein. Dies wird in Klammern vor dem Untertitel angezeigt. Geräusche werden entweder sachlich beschrieben (etwa in Deutschland) oder es werden Onomatopöien, also sprachliche Nachahmungen, verwendet (etwa in Schweden) (vgl. Jüngst 2010:133-137). Die

zusätzlichen Informationen sollen Atmosphäre vermitteln und besonders gehörlosen Menschen helfen, der Handlung besser folgen zu können (vgl. Díaz Cintas/ Remael 2014:14).

Bei Musik mit Text wird der Liedtext untertitelt. Entweder wird mit einer kleinen Note oder mit einer Raute am Anfang des Untertitels verdeutlicht, dass jemand singt. Falls der Liedtext keine Rolle spielt, wird ein Untertitel mit „Musik“ eingeblendet, der oft noch durch Adjektive spezifiziert wird, um anzuzeigen, welche Stimmung ausgedrückt werden soll (etwa fröhliche oder traurige Musik etc.) (vgl. Jüngst 2010:137).

Zudem sollten Untertitel so weit wie möglich die Vokalfolge im Bezug zum Originaltext widerspiegeln. Denn Zuschauer\*innen könnten an der Mundstellung von Sprecher\*innen erkennen, ob sich Originaltext und Untertitel stark voneinander unterscheiden. Außerdem ist es wichtig, dass die Untertitel möglichst synchron zum Originaltext gesendet werden. Ansonsten könnte dies zu Verwirrungen bei Menschen mit Resthörvermögen führen, weil sich ein Widerspruch zwischen Untertiteln und Dialog ergibt (vgl. Jüngst 2010:128). Dies ist bei simultan erstellten Live-Untertiteln jedoch nicht möglich. Anhand von vorbereiteten Untertiteln, die im semi-live Modus geschickt werden, kann jedoch eine mehr oder weniger genaue Synchronität geschaffen werden.

Je nach Textlänge und/oder Sprechgeschwindigkeit müssen Untertitel dementsprechend gekürzt werden. Denn Standzeit und Zeilenlänge sollen möglichst eingehalten werden, damit die Untertitel synchron zum Originaltext stehen. Laut Umfrage einer Arbeitsgruppe (vgl. Hezel 2009:232f.) sollte die Originalsprache jedoch weitestgehend erhalten bleiben und unbekannte Wörter sollten nur dann durch Synonyme ersetzt werden, falls die Sprechgeschwindigkeit das Lesetempo überschreitet. Da ein Teil der Zielgruppe noch ein gewisses Resthörvermögen hat oder Lippenlesen kann, nehmen Menschen mit Hörbeeinträchtigung es oft als Zensur oder Bevormundung wahr, falls die Abweichung von der Originalsprache zu groß ist.

Eine weitere Problematik stellt sich, wenn beim Originalton weder gesprochen wird noch Geräusche zu hören sind. Gehörlose Zuseher\*innen könnten sich über die fehlenden Untertitel wundern und glauben, dass diese ausgefallen sind (vgl. Jüngst 2010:128). Ausfälle kommen im Fernsehen ab und zu vor, vor allem bei der Live-Untertitelung ist das keine Seltenheit. Dies wird auch von den Zuschauer\*innen kritisiert (vgl. Hezel 2009:158f.). Um diese Problematik zu vermeiden, wird bei einigen Sendern mithilfe von „Überbrückungsuntertiteln“ explizit gemacht, dass es sich um stumme Szenen handelt (vgl. Hezel 2009:202). Näheres zur Live-Untertitelung folgt im nächsten Kapitel.



## **2. Live-Untertitelung**

Bei der Live-Untertitelung handelt es sich um einen Sonderfall der Untertitelung. Im vorliegenden Kapitel soll zuerst auf die Definition der Live-Untertitel und auf Allgemeines eingegangen werden. Im Anschluss wird kurz die Geschichte der Live-Untertitelung beleuchtet und danach werden die unterschiedlichen Methoden vorgestellt, wie Live-Untertitel hergestellt werden. Außerdem werden noch die Eigenheiten und Problematiken beim Live-Untertiteln beschrieben.

### **2.1. Definition und Allgemeines**

Im Fernsehen werden nicht nur im Voraus aufgezeichnete Sendungen gezeigt, sondern viele Programme werden live übertragen, so wie etwa Nachrichtensendungen, große Shows oder Sportereignisse. Das Interesse daran ist quer durch verschiedene Zielgruppen groß und somit ist es auch wichtig, dass alle Menschen Zugang dazu haben. Live-Untertitel leisten hierbei einen wichtigen Beitrag. Sie sind vor allem bei den öffentlich-rechtlichen Sendeanstalten zu finden und werden entweder von den Funkhäusern selbst erzeugt oder von professionellen Untertitelungsfirmen (vgl. Jüngst 2010:138). Mittlerweile werden auch bei privaten Sendern immer mehr Live-Sendungen untertitelt. Zwar wünscht sich die Zielgruppe mehr Live-Untertitel, allerdings wurde – zumindest kurz nach deren Einführung – die Qualität in Deutschland als mangelhaft angesehen. Viele standen den semi-live Untertiteln eher kritisch gegenüber, weil sie in den Augen der Zuseher\*innen keine „echten“ Live-Untertitel seien. Außerdem kamen Auslassungen und verzögerte Untertitel zu häufig vor (vgl. Hezel 2009:158ff.).

Die Live-Untertitelung ist ein noch relativ neues Feld in der Translationswissenschaft. Sie findet aber mittlerweile ihren Platz bei Fachkongressen, wo über Probleme und Facetten diskutiert wird. Ein Forschungsprojekt ist beispielsweise das ILSA-Projekt zur interlingualen Live-Untertitelung, das von der Europäischen Union von 2017 bis 2020 gefördert wurde (vgl. ILSA o.J.). Bei Live-Untertiteln (auch „real-time subtitles“ (Romero-Fresco 2011:6)) handelt es sich um eine spezielle Art von Untertiteln für SDH-Publikum, die in Live-Sendungen verwendet wird. Sie findet meist intralingual statt, allerdings werden auch interlinguale Live-Untertitel in letzter Zeit immer öfters produziert (vgl. Romero-Fresco 2011:13). Die Live-Untertitler\*innen versuchen, den Kontext einer Sendung im Vorhinein zu verstehen, und schicken die Untertitel dann in Echtzeit ab (vgl. EBU 2004:10). Dazu muss hinzugefügt werden, dass Live-Untertitel nicht nur Live-Sendungen betreffen, sondern auch kurz vorab aufgezeichnete Sendungen, wobei die Untertitler\*innen allerdings keine Zeit hatten, die Untertitel im Voraus

zu erstellen (vgl. Romero-Fresco 2011:1). Romero-Fresco (2011:12) unterscheidet hierbei drei Arten:

- Live: Diese Sendungen werden in Echtzeit ausgestrahlt, wie zum Beispiel Nachrichten oder Sportveranstaltungen.
- As-Live: Dabei handelt es sich um Sendungen, die mit kurzer Verzögerung ausgestrahlt werden, um eventuell unerwünschtes Material nicht auf Sendung zu schicken.
- Voraufzeichnung: Ein Großteil der Sendungen wird vorab aufgezeichnet und vor der Übertragung bearbeitet.

Somit bezieht sich der Begriff „live“ beim Untertiteln nicht auf die Sendung selbst, sondern auf den Zeitpunkt, wann die Untertitel erstellt bzw. gesendet werden, wie schon in Kapitel 1.6.2 erläutert wurde. Live-Untertitel werden demnach entweder simultan während der Sendung produziert oder im semi-live Modus geschickt (vgl. Jüngst 2010:138). Das „hybride Verfahren“ (Jüngst 2010:141) bezeichnet jene Live-Sendungen, bei denen semi-live Untertitel und simultan erstellte Untertitel abwechselnd gesendet werden<sup>4</sup>. Das kann man zum Beispiel bei Nachrichtensendungen finden, bei denen Videos schon geschnitten und mit Text versehen sind. Die Live-Untertitler\*innen können diese Texte meist im Vorhinein vorbereiten und im semi-live Modus senden. Jedoch gibt es keine hundertprozentige Sicherheit, dass die Beiträge und Texte nicht kurz vor Sendung verändert oder neu geschnitten werden, oder sogar ganz ausfallen oder neue hinzukommen, wie etwa bei Eilmeldungen (vgl. Jüngst 2010:141; Romero-Fresco 2011:12; Lindner 2016:312). Aufgrund der hohen Konzentration, die Live-Untertitler\*innen an den Tag legen müssen, gilt die Live-Untertitelung als kognitiv anstrengendste Form der Untertitelung (vgl. Jüngst 2010:138f.).

Um ein professionelles Ergebnis ohne gravierende Fehler abliefern zu können, liegt bei wichtigen Ereignissen der Text meist in voller Länge im Vorhinein vor. Somit können sich die Untertitler\*innen dementsprechend vorbereiten – sofern der Text auch für sie verfügbar ist. Hierbei handelt es sich oft um sehr komplexe Texte, die für simultan erstellte Untertitel wenig geeignet wären. Auch sonst können Untertitler\*innen im Normalfall auf eine Datenbank zugreifen und sich vor Sendungsbeginn mit den wichtigsten Informationen vertraut machen (vgl. Jüngst 2010:141).

---

<sup>4</sup> Im weiteren Verlauf der Arbeit wird der Begriff „gemischte Untertitel“ verwendet.

## 2.2. Geschichte und Methoden der Live-Untertitelung

Im Laufe der Zeit entwickelten sich verschiedene Methoden für die Erstellung von Live-Untertiteln. Zum einen produzierten Stenograf\*innen die Untertitel mittels speziell entwickelter Tastaturen und Techniken. Dies ähnelt der Stenografie-Arbeit, die in Gerichten benutzt wird (vgl. Jüngst 2010:139; Robson 2004:119-148, zit. nach Díaz Cintas/ Remael 2014:20). Zum anderen wurden Live-Untertitel mithilfe einer Spracherkennungssoftware erstellt. Hierbei diktieren die Untertitler\*innen die Untertitel in ein Mikrofon, das an einen Computer angeschlossen ist, und ein Spracherkennungsprogramm wandelt das gesprochene Wort in einen schriftlichen Text um.

Die ersten Live-Untertitel gab es in den USA im Jahre 1982. Gerichtsstenograf\*innen bekamen ein spezielles Training und erstellten anhand von Stenografiermaschinen Untertitel für gehörlose Menschen und Menschen mit Hörbeeinträchtigung. Das Respeaking (siehe Kapitel 3) fand seinen Ursprung ebenfalls in den USA mit dem Gerichtsstenografen Horace Webb. Dabei handelt es sich um eine besondere Methode für die Erstellung von Live-Untertiteln mithilfe von Spracherkennungsprogrammen. Um Gerichtsverhandlungen nicht anhand von Notizen transkribieren zu müssen, unternahm er in den 1940er Jahren erste Versuche, indem er den Text über ein Mikrofon aufnahm, um ihn anschließend zu transkribieren. Aufgrund der Lautstärke und überlappender Äußerungen war dies aber nicht möglich. Webb schlug deswegen vor, dass Gerichtsstenograf\*innen jedes Wort der Originalrede mithilfe einer Stenomaske, die die Lautstärke beim Sprechen minimierte, in ein Mikrofon diktieren. Das Diktierte wurde dann zur Erstellung der Transkription benutzt. Zwar wurde bei dieser Methode keine Spracherkennungssoftware verwendet – und somit gab es auch keinen simultan erstellten Text –, allerdings wurde das Grundprinzip des Respeakings angewandt. Dies kann somit als Vorstufe des Respeakings angesehen werden (vgl. Romero-Fresco 2011:7). Zwar gab es diese Technologie schon seit den 1990er Jahren, allerdings konnte sie sich aufgrund der relativ häufig auftretenden Schreib- und Erkennungsfehler für lange Zeit nicht gegen die herkömmlichen Methoden zur Erstellung von Live-Untertiteln durchsetzen. Zudem boten die Spracherkennungsprogramme am Anfang nur recht wenige Sprachen an. Es war aber nur eine Frage der Zeit, bis sie die herkömmlichen Methoden mit Tastatur ablösen würden (vgl. Jüngst 2010:139; Robson 2004:149-160, zit. nach Díaz Cintas/ Remael 2014:20). Das erste „realtime voice writing“ für Fernsehuntertitel gab es im Jahre 2003 (vgl. Romero-Fresco 2011:7). Heutzutage findet die Spracherkennung für die Erstellung von Live-Untertiteln in vielen Ländern immer mehr Anwendung.

In Europa hingegen gab es die ersten Live-Untertitel erst Ende der 1980er, Anfang der 1990er Jahre. Auch hier setzte man am Anfang auf die Methode der Stenografie, oder die Untertitel konnten im Vorhinein vorbereitet und im semi-live Modus gesendet werden (vgl. Romero-Fresco 2011:8f.). In Deutschland wurden die ersten Live-Untertitel im Fernsehen im Jahr 1985 bei einem Fußballspiel eingesetzt (vgl. Jüngst 2010:139). 1987 gab es Versuche mit dem Velotype (vgl. Nagel 2009:32), einer speziellen Tastatur aus den Niederlanden. Damit konnte die Schreibgeschwindigkeit deutlich erhöht werden (vgl. Lambourne 2006). Erste Tests mit dem Respeaking fanden in Spanien im Jahr 2004, in Italien 2008 statt (vgl. Romero-Fresco 2011:9).

Für die Erstellung von Live-Untertiteln benötigt man spezielle Techniken und Arbeitsgeräte. Auch aus diesem Grund gelten Live-Untertitel als noch recht junge Form der audiovisuellen Translation (vgl. Jüngst 2010:139). Auf die unterschiedlichen Methoden, die für die Produktion von Untertiteln verwendet werden, soll im Folgenden kurz eingegangen werden.

### **2.2.1. Eingabe über Tastaturen**

Bei der Eingabe über Tastaturen gibt es verschiedene Arbeitsweisen, die verwendet werden. Gearbeitet wird entweder mit einfachen, handelsüblichen Computertastaturen oder mit speziellen Tastaturen zum Schnellschreiben.

a) Mit gewöhnlichen QWERTZ-Tastaturen versuchen Live-Untertitler\*innen, so schnell wie möglich den gehörten Text zu verschriftlichen. Eine alte Form dieser Arbeitsweise besteht darin, dass der Text von einem/einer Redakteur\*in zusammenfasst und angesagt und von einer professionellen Schreibkraft direkt eingetippt wird. Diese Methode ist zwar effizient, aber durch den Einsatz von zwei Arbeitskräften auch relativ teuer (vgl. Jüngst 2010:140). Studien zeigen zudem, dass mit QWERTZ-Tastaturen zwar eine sehr hohe Genauigkeit erzielt werden kann, aber die Geschwindigkeit für die Praxis zu langsam ist. Die raschesten Schnellschreiber\*innen können rund 120 Wörter pro Minute tippen. Üblicherweise schaffen professionelle Schreibkräfte eher nur zwischen 50 und 100 Wörter pro Minute. Beide Werte weichen jedoch stark von der durchschnittlichen Sprechgeschwindigkeit von Fernsehmoderator\*innen ab, denn diese liegt bei etwa 180 Wörtern pro Minute (vgl. Romero-Fresco 2011:13).

Um die Anschlagzahl beim Tippen zu erhöhen, können Schriftdolmetscher\*innen (mehr dazu Kapitel 3.1.) Programme benutzen, mit denen sie Kürzel oder Textmakros anlegen, um Wörter damit auslösen zu lassen. „st“ wird somit automatisch zu „(Student/in)“. Diese Methode

verwenden Schriftdolmetscher\*innen bei besonders häufigen, in der Regel langen Wörtern sowie Regieanweisungen (vgl. Platter 2015:28ff.; Witzel 2020:315). Witzel (2020:315) argumentiert aber, dass eine schnelle Schreibgeschwindigkeit alleine noch nichts über die Qualität aussagt. Denn selbst starke Komprimierungen können trotz geringerer Anschlagzahl gute Ergebnisse erzielen. Zudem ist diese Methode am kostengünstigsten, jedoch unterliegt sie physischen Grenzen.

b) Eine weitere Methode ist das Tandem- bzw. Dual-Keyboard-System. Geübte schaffen etwa 150 Wörter pro Minute. Hierbei tippen zwei Schnellschreiber\*innen an zwei verschiedenen Tastaturen abwechselnd Sätze oder vervollständigen sie gegenseitig (vgl. Romero-Fresco 2011:13). Der untereinander aufgeteilte Text wird dann abwechselnd in das laufende Bild geschickt. Diese Arbeitsweise erfordert sehr viel Konzentration und ist eher für kürzere Sendungen vorgesehen (vgl. Jüngst 2010:140f.), führt allerdings auch zu qualitativ recht hochwertigen Untertiteln (vgl. Lambourne et al. 2004:270). Manchmal arbeiten sogar bis zu fünf Untertitler\*innen gleichzeitig an einer Sendung. Nach einer Art „Ampel-Methode“ sind sie parallel einsatzbereit und sobald ihr Bildschirm ein grünes Licht anzeigt, tippen sie mit der Computertastatur das Gehörte ab, bis das Licht wieder rot wird (vgl. Orero 2006).

c) Um die Geschwindigkeit beim Schreiben weiter zu erhöhen, wurde das sogenannte Velotype, heutzutage besser bekannt als Veyboard, erfunden. Hierbei handelt es sich um eine Silbenakkordtastatur, bei der mehrere Tasten gleichzeitig gedrückt werden können. Somit ist es möglich, gleich ganze Silben und Wörter, statt nur einzelne Buchstaben zu tippen. Zwar erhöht sich damit die Schreibgeschwindigkeit im Vergleich zu herkömmlichen Tastaturen, allerdings ist sie mit rund 90 bis 120 Wörtern pro Minute nach intensiver Übung noch deutlich langsamer als normales Sprechtempo (vgl. Lambourne et al. 2004:270). Des Weiteren sind die Kosten aufgrund der langen Ausbildung relativ hoch. Trotzdem setzen auch heute noch Länder wie Schweden und die Niederlande auf diese Methode (vgl. Geens 2007:101).

d) Mit der Maschinenstenografie können Schriftdolmetscher\*innen fast Wort-für-Wort-Transkripte mit einer speziellen Tastatur, dem Stenoclient, erstellen. Damit können mehrere Tasten gleichzeitig gedrückt werden, um so ganze Silben, Wörter oder gar Sätze mit einer einzigen Handbewegung zu schreiben. Im Gegensatz zum Velotype fehlen jedoch einige Buchstaben des englischen Alphabetes beim Stenoclient. Buchstabenkombinationen ersetzen die fehlenden Zeichen (vgl. Romero-Fresco 2011:15). Schriftdolmetscher\*innen stenografieren den Ton und

eine Software wandelt die getippten Silben in ganze Sätze um. Dies ermöglicht die Verarbeitung von sehr großen Textmengen (vgl. Jüngst 2010:140f.).

Die Maschinenstenografie wird besonders häufig in Gerichtsverhandlungen eingesetzt, wo eine hohe Schreibrate sehr wichtig ist. Professionelle Schnellschreiber\*innen schaffen damit etwa 200 bis 250 Wörter pro Minute, wenn nicht sogar 300. Die Technik gilt somit als schnellste für das Live-Untertiteln. Sowohl die Geschwindigkeit als auch die Genauigkeit sind sehr hoch. Viele Rundfunkanstalten in Europa setzten daher lange Zeit auf diese Methode. Die Technik war ideal für etwaige Einsätze in Live-Sendungen wie für Interviews oder Debatten (vgl. Romero-Fresco 2011:15). Allerdings ist es möglich, dass die gesendete Menge an Untertiteln aufgrund der Schnelligkeit so hoch ist, dass Zuseher\*innen kaum ausreichend Zeit zum Lesen haben (vgl. Jüngst 2010:140). Des Weiteren ist das Erlernen der komplexen Technik mit einer langjährigen Schulung verbunden und dementsprechend teuer (vgl. Romero-Fresco 2011:15).

Da die vorgestellten Methoden für die Erstellung von Live-Untertiteln entweder zu langsam oder zu kostenintensiv waren, suchte man nach Alternativen – vor allem auch, weil öffentlich-rechtliche Rundfunkanstalten immer mehr ihrer Sendungen für gehörlose Menschen und Menschen mit Hörbeeinträchtigung zur Verfügung stellen mussten. Daher setzte man immer mehr auf Spracherkennungsprogramme, weil diese einige Vorteile mit sich bringen.

### **2.2.2. Spracherkennung**

Bei der Erstellung von Live-Untertiteln mit Spracherkennung sprechen Live-Untertitler\*innen den Text sowie die Interpunktion mithilfe eines Mikrofons ein. In Sekundenschnelle verschriftlicht eine Spracherkennungssoftware den eingesprochenen Text (vgl. Jüngst 2010:139). Die Software muss im Vorhinein ausreichend trainiert werden, um sich an die Stimme eines/einer bestimmten Sprecher\*in zu gewöhnen. Je besser sie trainiert ist, desto besser funktioniert sie (vgl. Remael 2014:128). Um die Spracherkennung verwenden zu können, braucht es den Aufbau eines Sprachmodells. Der Datenbank muss zusätzlich zu den schon vorhandenen Vokabeln Terminologie eingespeist werden (vgl. Jekat 2014:94).

Üblicherweise arbeiten Live-Untertitler\*innen in schalldichten Kabinen, weil die Spracherkennungsprogramme sehr empfindlich auf externe Geräusche sind. Bevor der Untertitel auf Sendung geschickt wird, können die Untertitler\*innen Schreib- und Erkennungsfehler manuell mittels Computertastatur korrigieren (vgl. Jüngst 2010:139).

Zwar hat die Software bei der Erkennung von Konjugationen, Deklinationen und Komposita noch Probleme, allerdings werden mithilfe der Korrekturen der Live-Untertitler\*innen bei der Methode mit Spracherkennung gute Ergebnisse erzielt (vgl. Witzel 2020:316f.). Somit darf man die Spracherkennung nicht nur getrennt betrachten, sondern immer im Zusammenspiel mit einer herkömmlichen Computertastatur, mit der die Live-Untertitler\*innen Wörter gegebenenfalls tippen können, bei denen sie schon im Vorhinein antizipieren können, dass die Spracherkennungssoftware möglicherweise Probleme bei der Erkennung haben könnte bzw. im Nachhinein falsch erkannte Wörter ausbessern können, sofern sie diese rechtzeitig ausfindig machen. Als „unvoreilhaft“ bezeichnet Käber (2016:304) jedoch die Eigenschaft der Spracherkennungssoftware, Diktiertes immer auszuwerfen, auch wenn sie es nicht richtig verstanden hat. Fehlerhafte Untertitel sind beim Senden somit kaum zu vermeiden.

Grundsätzlich unterscheidet man bei der Spracherkennungsmethode zwischen zwei verschiedenen Arbeitsweisen: dem Respeaking und der Untertitel-Produktion eines\*einer unabhängigen Reporters\*Reporterin. Beim Respeaking werden die wichtigsten Informationen des Originaltons nachgesprochen. Jüngst spricht dabei von einer Art „Untertitel-Dolmetscher“ (Jüngst 2010:139). Bei der anderen Vorgehensweise beobachten die Untertitler\*innen das Bild eigenständig und verfassen den Untertitel unabhängig vom Originalton. Wichtige Informationen können vom gesprochenen Wort aber auch übernommen werden (vgl. Jüngst 2010:139).

Zudem lassen sich bei der Methode Pros und Contras finden. Der Nachteil besteht darin, dass Spracherkennungsprogramme nur auf ihnen bekannte Personen eingestellt sind und diese somit am besten erkennen. Jede\*r Sprecher\*in braucht ein eigenes Profil. Trotzdem kann es immer wieder zu Erkennungsungenauigkeiten kommen, was folglich zu fehlerhaften Untertiteln führt, falls etwa aus Zeitgründen die Untertitel nicht korrigiert werden können. Vor allem Eigennamen bereiten den Spracherkennungssoftwares Schwierigkeiten. Nichtsdestotrotz bietet diese Untertitelungsmethode auch Vorteile. Erstens ist das Arbeiten mit solchen Programmen weniger anstrengend als mit Tastaturen (vgl. Jüngst 2010:140). Zweitens können Live-Untertitler\*innen beim Respeaking eine nahezu normale Sprachgeschwindigkeit erreichen. Des Weiteren ist die Ausbildung mit einer Trainingsphase von zwei- bis dreimonatiger Dauer im Vergleich zu den Tastatur-Methoden auch relativ kurz. Aus diesen Gründen zählt das Respeaking neben der Maschinenstenografie zu den am häufigsten verwendeten Methoden, um Live-Untertitel zu erstellen (vgl. Lambourne 2006).

### 2.2.3. Vergleich der Methoden zur Erstellung von Live-Untertiteln

Lambourne (2007, zit. nach Romero-Fresco 2011:15) vergleicht die oben vorgestellten Methoden auf verschiedene Faktoren: Verzögerung, Geschwindigkeit, Genauigkeit, Ausbildung, Kosten und Auftreten schwerwiegender Fehler. Die herkömmliche Computer-Tastatur wird hierbei nicht mehr mit einbezogen, weil diese für das Erstellen von Live-Untertiteln kaum mehr Verwendung findet.

	<b>Delay</b>	<b>Speed (wpm)</b>	<b>Accuracy</b>	<b>Training</b>	<b>Cost</b>	<b>Seriousness of errors</b>
<b>Velotype</b>	Medium	Medium (90-120)	95%	12 months	Medium	High
<b>Dual keyb</b>	Medium	Medium/High 140-150	95-98%	6 months	Medium	Low
<b>Stenotype</b>	Low	Very high 220 – up to 300	97-98%	3 years	High	Medium
<b>Respeaking</b>	Low	High 160-190	97-98%	2-3 months	Low	Medium-High

Abbildung 1: Vergleich der Methoden zur Erstellung von Live-Untertiteln (Lambourne 2007, zit. nach Romero-Fresco 2011:15)

Wie aus der Tabelle ersichtlich, bieten Stenotype und Respeaking die größten Vorteile für die Erstellung von Live-Untertiteln. Sie erfüllen besonders gut die wichtigen Kriterien des kurzen zeitlichen Verzuges von Originalton und Untertitel, der schnellen Geschwindigkeit sowie der hohen Genauigkeit. Die häufigsten Fehler sind kurze, ausgelassene Wörter, kleine fehlerhafte Wörter und teilweise falsch angezeigte Wörter. Das Spracherkennungssystem macht grundsätzlich von sich aus keine Rechtschreibfehler (es sei denn, der\*die Respeaker\*in hat sich beim Erlernen und Einspeisen von neuen Wörtern vertippt). Das größte Problem der Spracherkennung ist, dass Verwechslungsgefahr bei Homophonen besteht sowie die Grammatik fehlerhaft sein kann. Zwar treten beim Respeaking im Vergleich zum Stenotype schwerwiegendere Fehler auf, allerdings sprechen die deutlich kürzere Ausbildungsphase und die günstigere Erstellung wiederum für die Methode (vgl. Lambourne 2006). Um das Problem der Homophone zu umgehen, können sich Respeaker\*innen im Sprachprofil eigens dafür zugeschnittene Makros anlegen. Das Respeaking soll im folgenden Kapitel näher beleuchtet werden. Zuvor wird noch auf die besonderen Herausforderungen bei der Live-Untertitelung eingegangen.



### 2.3. Herausforderungen bei der Live-Untertitelung

Da den Respeaker\*innen oftmals die Beiträge nicht lange genug vor Sendungsbeginn vorliegen, sodass keine Bearbeitung im Vorfeld möglich ist, müssen sie die Untertitel simultan erstellen. Bei diesem Prozess kommt es zu einigen Anforderungen und Herausforderungen. Aufgrund der hohen Sprechgeschwindigkeit in den Originaltexten, der Décalage und der Merkmale des Transfers von gesprochener in geschriebene Sprache ist es oft notwendig, die Originaltexte zu kürzen und/oder umzuformulieren. Dies alles geschieht unter hohem Zeitdruck. Auch Textkorrekturen können parallel durchgeführt werden. Diese Faktoren können zu einer Informationsverschiebung zwischen Originaltext und Live-Untertiteln sowie zu einer Verzögerung der Untertitel führen (vgl. Jekat 2014:88).

Live-Untertitel können entweder wörtlich oder zusammengefasst sein. Dies zählt zu einem der größten Streitpunkte in der SDH-Zielgruppe. Es gibt unterschiedliche Standpunkte dazu: Einerseits bevorzugen Sendeanstalten, Unternehmen und Gehörlosenbünde wörtliche, und somit schnellere Live-Untertitel. Andererseits sprechen sich Forscher\*innen für bearbeitete und demnach langsamere Live-Untertitel aus (vgl. Romero-Fresco 2009). Da nur die Methode der Stenografie mit der Sprechgeschwindigkeitsrate bei Fernsehsendungen mithalten kann, um wortgetreu Gesprochenes in Geschriebenes umzuwandeln, ist die häufigste Praxis das fast-wörtliche Untertiteln („near-verbatim subtitling“). Untertitler\*innen versuchen dabei so wortgetreu wie möglich zu sein. Jedoch gilt es als kaum machbar, den Originalton eins zu eins in der Live-Situation wiederzugeben (vgl. Romero-Fresco 2011:17). Diese Ansicht teilt auch Lindner (2016:308), weil zudem Menschen mehr Zeit zum Lesen als zum Hören brauchen. Trotzdem betont er, dass die Live-Untertitler\*innen bei allen Sendungsarten wie etwa bei Nachrichten, Dokumentationen oder fiktionalen Formaten darauf achten, keine relevanten Informationen auszulassen. Die Live-Untertitler\*innen entscheiden in der Situation selbst, was wichtig ist und was nicht. Normalerweise liefern Live-Untertitler\*innen auch Informationen über Geräusche (Jubel, Applaus) und färben Untertitel je nach Sprecher\*innen in unterschiedlichen Farben ein (Romero-Fresco 2011:17).

Die Herausforderungen beim Live-Untertiteln für nicht-geskriptete Texte können laut Lambourne (2006) folgendermaßen zusammengefasst werden:

- Erstellen von kohärenten und den Konventionen entsprechenden Untertiteln von mündlich dargebotenem, frei formuliertem Text mit einer der zuvor präsentierten Techniken. Dies beinhaltet die Verwendung richtiger Grammatik und Rechtschreibung, geringer Zeitverzögerung

von gesprochenem Text und Aufscheinen der Untertitel, inhaltliche Vollständigkeit sowie sachliche und typographische Richtigkeit.

- Die Untertitel so zu präsentieren, dass sie einfach gelesen bzw. verstanden werden können. Dies betrifft sowohl die Art der Untertitel-Präsentation (etwa passende Sinneinheiten, Pyramidenform (vgl. Hezel 2009:213-216; Gottlieb 2002:208)) als auch der zeitliche Abstand zum Original.

- Positionierung der Untertitel, damit relevante visuelle Informationen wie Inserts von Namen nicht verdeckt werden.

- Vermeidung von Programmverzögerungen.

Um die Herausforderungen beim Live-Untertiteln so gering wie möglich zu halten, werden die Sendeanstalten ersucht, zur Verbesserung und Bewältigung etwaiger Probleme beizutragen, indem sie zum Beispiel vorgeskriptetes Material, voraufgezeichnete Videos und Interviews vor Sendungsbeginn zur Verfügung stellen. Für Material, das vor – oder gar während – der Sendung abgerufen und transkribiert werden kann, können die Live-Untertitler\*innen die Untertitel vorbereiten und eventuell zeitlich vorprogrammieren. Dementsprechend sind vorbereitete Untertitel weniger fehlerhaft als reine live-produzierte Untertitel und der zeitliche Abstand der Untertitel zum Original bei Sendungsausstrahlung ist geringer. Im Endeffekt bleibt nur noch der „echte“ Live-Text übrig, der mit den zuvor beschriebenen Methoden untertitelt wird (vgl. Lambourne 2006).

Je mehr Untertitel im Vorhinein vorbereitet werden können, desto weniger Untertitel müssen live produziert werden, was sich wiederum auf die Untertitel-Qualität allgemein auswirkt. Es kann davon ausgegangen werden, dass die kognitive Leistung der Live-Untertitler\*innen abnimmt, je länger der reine Respeaking-Teil andauert. Diese Annahme bekräftigt auch Lambourne (2006) und führt an, dass aus diesem Grund die Live-Untertitler\*innen bei gewissen Sendeformaten mit längerer Dauer, wie zum Beispiel bei politischen Live-Talkshows, im Team arbeiten, um die Arbeitslast auf mehrere Personen aufzuteilen. Jedoch sind Fehler bei der Live-Untertitelung fast unvermeidbar und eine hundertprozentige Genauigkeit kann kaum erreicht werden (vgl. Lambourne 2006). Wie die Arbeitsweise beim Respeaking aussieht, wird im nachfolgenden Kapitel behandelt.

### 3. Respeaking

In diesem Kapitel wird näher auf das Respeaking und die Arbeitsweise eingegangen. Zudem wird ein Vergleich zwischen Respeaking und Simultandolmetschen vorgestellt. Des Weiteren werden kurz die Ausbildung und Kompetenzen von Respeaker\*innen beleuchtet. Anschließend wird das Effort Model nach Gile (2009) beschrieben und auf das Respeaking ausgeweitet, um die problemauslösenden Faktoren (*problem triggers*) sowie Bewältigungsstrategien (*coping tactics*) beim Respeaking aufzuzeigen. Zudem werden problemauslösende Faktoren beim Respeaking und aktuelle Untersuchungen zum Respeaking in der Live-Untertitelung dargelegt.

#### 3.1. Definition – Respeaking und Schriftdolmetschen

Unter Respeaking versteht Romero-Fresco (2011:1) die Erstellung von (meist intralingualen) Untertiteln mit Hilfe von Spracherkennung. Beim Respeaking handelt es sich um ein noch recht neues, unerforschtes Feld in der Wissenschaft. Somit gibt es in vielen Sprachen keine durchgängige Terminologie (vgl. Romero-Fresco 2011:2). Daher lassen sich auch im Englischen einige unterschiedliche Bezeichnungen finden wie

[...] *speech-based live subtitling* (Lambourne et al. 2004), (real time) *speech recognition-based subtitling* and *real-time subtitling via speech recognition* (Eugeni 2008), [...] *speech captioning* [...] or *shadow speaking* (Boulianne et al. 2009), [...] *revoicing* (Muzii 2006), *voice-writing* (Vincent 2007) and *realtime voice writing* (Keyes 2007) [...]. (Romero-Fresco 2011:2; Hervorhebungen und Kursivsetzung im Original)

Der Begriff Respeaking hat sich mittlerweile sowohl in der Berufswelt als auch in der Wissenschaft etabliert. Im Deutschen wird die Lehnübersetzung Respeaking verwendet (vgl. Romero-Fresco 2011:2f.). Die Aufgabe besteht für die Respeaker\*innen darin, den Sinn des Ausgangstextes zu verstehen und ihn schriftlich in Form von Untertiteln wiederzugeben (vgl. Romero-Fresco 2011:128). Die Methode wird allerdings nicht nur für die Erstellung von Untertiteln für zahlreiche Fernsehsendungen verwendet, sondern etwa auch für Veranstaltungen wie etwa in Museen, bei Konferenzen, Gesprächsrunden, bei universitären Lehrveranstaltungen, in der Schule, bei Geschäftstreffen oder Telefongesprächen (vgl. Romero-Fresco 2011:1f.; Romero-Fresco 2019:100).

Das Respeaking lässt sich in den Bereich des Schriftdolmetschens einordnen. Beim Schriftdolmetschen handelt es sich um die simultane Verschriftlichung von lautsprachlichen und paraverbalen Äußerungen. Die Tätigkeit ist eine Kommunikationshilfe für gehörlose Menschen und Menschen mit Hörbeeinträchtigung (vgl. Tiittula 2006:481f.), um ihnen die Teilhabe in Live-Situationen zu ermöglichen (vgl. Mälzer/ Wünsche 2020:331). Das Schriftdolmetschen

weicht vom Live-Untertiteln dahingehend ab, dass nur die gesprochenen Äußerungen angezeigt werden und die charakteristischen Eigenheiten von Untertiteln nicht gelten. Denn zu Untertiteln kommen noch zahlreiche andere Faktoren hinzu, wie etwa das begrenzte Zeichenlimit oder dass ein Untertitel aus höchstens zwei Zeilen bestehen soll (vgl. Romero-Fresco 2011:1f.; Romero-Fresco 2019:100). Der Zieltext erscheint beim Schriftdolmetschen in Wörtern mehr oder weniger zeitgleich mit dem Ausgangstext. Im Gegensatz dazu erscheinen Untertitel auf dem Bildschirm üblicherweise als fertige Textblöcke (vgl. Tiittula 2006:481f.).

## **3.2. Arbeitsweise**

Zwar ist es schwierig, eine genaue Beschreibung der Arbeitsweise beim Respeaking zu bieten, weil sie entweder zu vereinfacht oder zu komplex wäre, aber grob gesagt besteht die Technik darin, dass die Respeaker\*innen den Originalton einer Sendung hören und den Text inklusive Satzzeichen nachsprechen. Eine Spracherkennungssoftware wandelt die Äußerungen mit der kleinstmöglichen Verzögerung in Untertitel um. Romero-Fresco stellt jedoch klar, dass es sich bei Respeaking nicht nur um einfaches Wiederholen des gehörten Textes handelt. Denn aufgrund der Schnelligkeit des Ausgangstextes können die Respeaker\*innen die Wörter eines Textes nicht einfach nachsprechen, sondern müssen sie umformulieren, um möglichst synchron zum Original zu bleiben (vgl. Romero-Fresco 2011:1). Außerdem ist die Redegeschwindigkeit höher als die Lesegeschwindigkeit. Diese Faktoren führen dazu, dass die Respeaker\*innen den Ausgangstext paraphrasieren: “[T]he respeaker [...] has to perform a process of message comprehension and reformulation that often requires a certain distancing from mere word-for-word formulation.“ (Romero-Fresco 2011:47) Auch Luyckx et al. (2010) argumentieren, dass beim Live-Untertiteln eine wörtliche Wiedergabe – auch wenn diese vom Zielpublikum gewünscht ist – kaum umsetzbar und eine Reduzierung des Ausgangstextes daher unvermeidbar ist. Allerdings handelt es sich hierbei nicht um willkürliche Auslassungen, sondern um Textstellen mit nachrangigen oder überflüssigen Informationen. Somit hat dies keine Auswirkungen auf das Verstehen des Ausgangstextes.

### **3.2.1. Respeaking-Prozess bei der Live-Untertitelung**

Um die einzelnen Schritte während des Respeakings aufzuzeigen, entwerfen Luyckx et al. (2010:4) eine zeitliche Darstellung (siehe Abb. 2) für den Respeaking-Prozess bei der Live-Untertitelung.

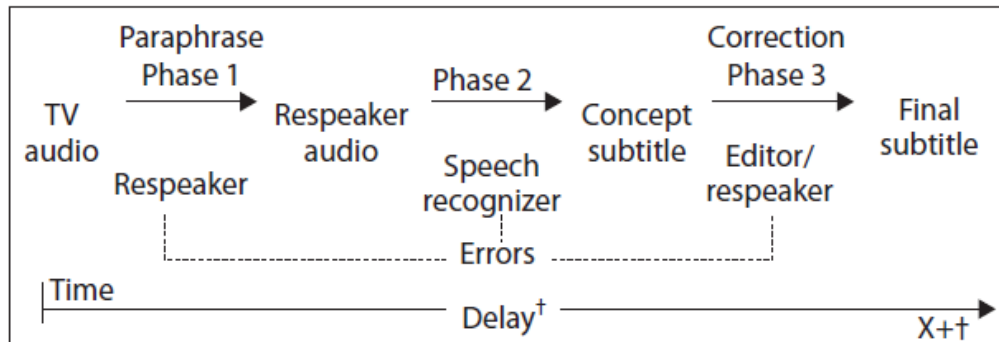


Abbildung 2: Zeitliche Darstellung des Respeaking-Prozesses bei der Live-Untertitelung (Luyckx et al. 2010:4)

Ausgehend von einem Audiosignal der zu untertitelnden Sendung (TV audio) hört der/die Respeaker\*in den Ausgangstext und wiederholt bzw. paraphrasiert das Gesagte (Respeaker audio). Eine Software transkribiert den Text (Concept subtitle) in ein Textfeld. Etwaige Fehler können von dem\*der Respeaker\*in selbst oder von ihren Kolleg\*innen ausgebessert werden, bevor der fertige Untertitel (Final subtitle) gesendet wird (vgl. Luyckx et al. 2010:4). Oft sind Respeaker\*innen für eine oder mehrere Sendungen zuständig. Arbeiten sie im Team, was üblicherweise bei längeren Sendungen der Fall ist, bereiten beide Mitarbeiter\*innen die Inhalte vor. Während der Live-Sendung sendet eine\*r von ihnen die vorbereiteten Untertitel synchron zum Bild, die andere Person ist für das Respeaking zuständig (vgl. Lindner 2016:307f.). Es ist aber auch gängig, dass in dieser Aufgabenverteilung zwischen beiden Kolleg\*innen abgewechselt wird.

### 3.2.2. Arbeitswerkzeuge

Respeaker\*innen verwenden einen PC als Arbeitsgerät und tragen Kopfhörer, um den Originalton der Sendung zu hören. Zudem kommt ein USB-Mikrofon zur Anwendung. Auf einem Fernsehbildschirm können sie überprüfen, ob die Untertitel auch tatsächlich gesendet werden (vgl. Romero-Fresco 2019:98).

Für das Erstellen der Live-Untertitel verwenden Respeaker\*innen zwei verschiedene Softwarearten: Zum einen nutzen sie eine Spracherkennung (auch mit SR abgekürzt für speech recognition), die das gesprochene Wort verschriftlicht. Zum anderen benutzen sie eine Untertitelungssoftware, in der die verschriftlichten Äußerungen als Untertitel am Bildschirm eingeblendet werden (vgl. Romero-Fresco 2011:1f.).

Von der Spracherkennungssoftware hängt zudem ab, in welcher Form die Untertitel am Bildschirm aufscheinen. Entweder erscheint der Text blockartig (etwa bei Dragon Naturally Speaking) in ein-, zwei- oder dreiteiligen Zeilen, in Satzteilen oder ganzen Sätzen. Oder er wird fortlaufend angezeigt (wie bei ViaVoice), wo das Gesprochene Buchstabe für Buchstabe, Silbe

für Silbe oder (meistens) Wort für Wort angezeigt wird. Wie die Untertitel aufscheinen, können Respeaker\*innen je nach deren Präferenz selbst bei der Untertitelungssoftware einstellen. Oft liegt letztendlich die Entscheidung bei den Sendeanstalten oder Untertitelungsfirmen, die verschiedene Faktoren wie Zeitverzug, Genauigkeit und Lesefreundlichkeit berücksichtigen (vgl. Romero-Fresco 2011:17). Romero-Fresco (2010) argumentiert, dass blockartig aufscheinende Untertitel besser geeignet sind, weil die Rezipient\*innen sich dann deutlich länger auf das Bild konzentrieren können und sich nicht so sehr auf die Untertitel fixieren müssen.

Des Weiteren benutzen Respeaker\*innen ein Programm, mit dem sie Zugriff auf die Sendungsinhalte und teilweise auch Texte haben (vgl. Romero-Fresco 2019:98f.).

### **3.2.3. Workflow**

Als Vorbereitungsmaterial dienen etwa Manuskripte, Dialogbücher, Moderationstexte und Bildmaterial. Für bestimmte Sendungen wie Nachrichten- oder Magazinsendungen lassen sich Informationen zum Sendungsablauf, Beitrags- und Moderationstexte, Namen der Gäste und Videos auf einem Redaktionssystem (etwa Open Media) sowie auf Video-Produktionssystemen finden und für die Vorbereitung nutzen. Bestenfalls sind alle Inhalte der Sendung verfügbar und die Respeaker\*innen können diese gemäß der Untertitel-Richtlinien bzw. -Standards vorbereiten. Allerdings sind kurzfristige Änderungen immer möglich, wie etwa bei Sendungsabläufen oder neuen Beiträgen bei aktuellen Nachrichtenformaten. Dadurch kann eine vollständige Vorbereitung nicht immer gegeben sein (vgl. Lindner 2016:308f.). Respeaker\*innen müssen daher ein hohes Maß an Flexibilität und Konzentration aufbringen (vgl. Lindner 2016:312).

Alle anderen Teile, die nicht vorbereitet werden konnten, weil kein Text zur Verfügung stand oder es zu Änderungen kam, untertiteln die Respeaker\*innen live mittels der Respeaking-Methode (vgl. Romero-Fresco 2019:98f.). Wie schon erwähnt, wiederholen sie den Text nicht wortwörtlich, sondern müssen ihn anpassen. Da die Datenbank der Spracherkennungssoftwares nicht alle Wörter beinhaltet und somit vor allem Termini und Eigennamen oftmals nicht erkannt werden, bedarf es einer Anpassung des Textes. Die Respeaker\*innen müssen daher versuchen, nicht vorhandenes Vokabular zu vermeiden. Folglich ist es auch sinnvoll, unbekannte Wörter im Vorhinein einzuspeichern, sofern Vorbereitungsmaterial verfügbar ist (Romero-Fresco 2011:47f.).

Für die Erstellung von Live-Untertiteln gilt es zudem, den zeitlichen Abstand zwischen Originalton und Untertiteln so kurz wie möglich zu halten. Ein guter Richtwert wäre etwa ein Abstand von zwei bis drei Sekunden (vgl. Geens 2007:103ff.). Die Verzögerung hängt von

vielen verschiedenen Faktoren ab wie etwa der Software, der Korrekturmethode oder der Untertitelungsart. Der Verzug bei verschiedenen Softwareprogrammen kann unterschiedlich lang sein und es kommt auch darauf an, wie viele Personen für die Untertitel gleichzeitig zuständig sind (vgl. Romero-Fresco 2011:2).

Nach Sendungsende überarbeiten die Respeaker\*innen die Untertiteldateien für mögliche Wiederholungen, die Weiterverwertung in Mediatheken oder für den Programmaustausch mit anderen Fernsehanstalten (vgl. Lindner 2016:309).

#### **3.2.4. Korrekturmethoden**

Da es etwa aufgrund der Spracherkennungssoftware immer wieder zu Fehlern beim Respeaking kommt, müssen sich die Live-Untertitler\*innen entscheiden, welche Korrekturmethode sie anwenden, um Fehler auszubessern. Romero-Fresco (2011:16) unterscheidet zwischen drei Methoden: keine Korrektur, Selbstkorrektur und parallele Korrektur. Dies hängt davon ab, ob das Augenmerk darauf liegt, fehlerfreie Untertitel zu senden oder den Zeitabstand zwischen Original und Untertitel möglichst gering zu halten. Nicht zu korrigieren kommt allerdings sehr selten vor, auch wenn so der zeitliche Verzug am geringsten wäre.

Die am häufigsten angewandte Methode ist jene der Selbstkorrektur. Die Respeaker\*innen korrigieren ihre eigenen Fehler, nachdem der Untertitel schon auf Sendung gegangen ist. Zwar variieren die Lösungsstrategien von Land zu Land bzw. Unternehmen zu Unternehmen, aber meistens wird nach dem Fehler das richtige Wort, das einem Bindestrich folgt, geschickt. Obwohl es hier zu einem zeitlichen Verzug kommt, ist dieser kleiner als bei der parallelen Korrekturmethode. Allerdings ist die Selbstkorrektur ein weiterer Stressfaktor zur ohnehin schon anspruchsvollen Erstellung von Live-Untertiteln (vgl. Romero-Fresco 2011:16). Im Normalfall können die Respeaker\*innen die Untertitel auch noch vor dem Senden korrigieren.

Die Fehler treten aus unterschiedlichen Gründen auf: Die Respeaker\*innen könnten den Ausgangstext oder Teile davon missverstanden, falsch aufgefasst oder Wörter undeutlich ausgesprochen haben, was zur Folge hat, dass die Spracherkennungssoftware etwas anderes als gewollt erfasst. Im Anschluss könnten die Respeaker\*innen den Fehler nicht erkannt und somit nicht ausgebessert haben. Es könnte gar der Fall sein, wenn dies aber eher selten vorkommt, dass die Respeaker\*innen beim Ausbessern selbst noch Fehler – etwa durch das (Ver-)Tippen mit der Tastatur – hinzufügen (vgl. Remael et al. 2014:124f.).

Das Korrigieren geschieht – je nach Anzahl der zusammenarbeitenden Kolleg\*innen sowie je nach Schwierigkeit und Länge des Programms – entweder alleine („mono-LS model“) oder mithilfe eines\*einer zweiten Respeakers\*Respeakerin („duo-LS model“). Die erste Person

konzentriert sich auf das Respeaking an sich, und bessert, wenn überhaupt, nur kleine Fehler selbst aus. Die zweite Person kümmert sich generell um die Korrekturen. Als dritte Variante kann noch das „multi-LS model“ angeführt werden, wo Sendungen nicht direkt live, sondern mit etwas Verzug ausgestrahlt werden. So könnte man die Qualität der Untertitel verbessern und den Abstand des Gesagten und der eingeblendeten Untertitel verkürzen (vgl. Remael et al. 2014:124-127). Bei der parallelen Korrektur korrigieren ein\*e oder gar zwei Kolleg\*innen die Untertitel, noch bevor diese auf den Fernsehbildschirmen zu lesen sind. Dies führt zwar zu (fast) fehlerfreien Untertiteln, allerdings ist die Verzögerung hier auch mit am größten (vgl. Romero-Fresco 2011:16).

Aufgrund der Überlappungen verschiedener Leistungen beim Respeaking sollen im nächsten Unterkapitel das Respeaking und das Simultandolmetschen gegenübergestellt werden.

### **3.3. Respeaking und Simultandolmetschen – eine Gegenüberstellung**

Um beide Tätigkeiten miteinander vergleichen zu können, soll zuerst eine kurze Definition des Simultandolmetschens gegeben werden. Der internationale Verband der Konferenzdolmetscher\*innen (AIIC) definiert das Simultandolmetschen wie folgt:

Das gesprochene Wort wird zeitgleich – also simultan – in eine oder beliebig viele andere Sprachen übersetzt. Dies erfordert vom Dolmetscher höchste Konzentrationsfähigkeit. Die Dolmetscher wechseln sich deshalb alle 20 bis 30 Minuten ab. Beim Simultandolmetschen arbeiten die Konferenzdolmetscher in schallisolierten Kabinen, die entweder im Saal fest eingebaut sind oder als transportable Anlage für die Veranstaltung angemietet und aufgebaut werden. Die Dolmetscher hören den Redner über Kopfhörer und sprechen über Mikrofon in einen Übertragungskanal. (AIIC 2022)

Eugeni (2010) sieht durch das gleichzeitige Zuhören und Reden eine ähnliche Leistung beim Respeaking und Simultandolmetschen:

The characteristic that more resembles the two professionals is the skill to carry out several cognitive activities simultaneously in order to create an aural output that is as faithful as possible to the source text. More specifically, what the two professionals share is the ability to listen, analyze and speak at the same time. (Eugeni 2010:2)

Zudem hebt er hervor, dass es bei beiden Tätigkeiten darum geht, eine Ausgangsrede für ein Zielpublikum verständlich zu machen, das ohne die Hilfe der Dolmetscher\*innen bzw. Respeaker\*innen die Botschaft nicht verstehen könnte. Den Prozess und die Funktion der Dolmetschungen beider Arten sieht er daher als sehr ähnlich oder gar identisch an. Der Prozess ist das Ergebnis einer sprachlichen Reformulierung: beim Simultandolmetschen die sprachliche Reformulierung von einer Sprache in die andere; beim Respeaking die Reformulierung einer sprachlichen Struktur in eine andere – einfachere, kürzere und entsprechend den Eigenschaften



der Schriftsprache. Als Funktion führt er an, dass polysemiotische Texte einem Teil des Publikums zugänglich gemacht werden, welches ohne Verdolmetschung den Ausgangstext nicht verstehen würde (vgl. Eugeni 2010:2f.).

Auch Romero-Fresco (2019) argumentiert, dass das Respeaking als eine Art von Dolmetschen betrachtet werden kann: „It is, in effect, a form of (usually intralingual) computer-aided simultaneous interpreting with the addition of punctuation marks and features such as the identification of different speakers” (Romero-Fresco 2019:96). Zwar sind die meisten Untertitel, die durch Respeaking erstellt werden, intralingual, allerdings können sie auch interlingual sein. Somit kann das Respeaking laut Szarkowska et al. (2019) generell als eine Form von intra- oder interlingualer Translation betrachtet werden.

Im Folgenden soll näher auf die Gemeinsamkeiten und Unterschiede von Respeaking und Simultandolmetschen eingegangen werden. Da das Schriftdolmetschen (mit Spracherkennung) als Überkategorie vom Respeaking betrachtet werden kann, wo jedoch die charakteristischen Eigenheiten von Untertiteln nicht gelten, wird dieses für die Gegenüberstellung mit dem Simultandolmetschen herangezogen.

### **3.3.1. Gemeinsamkeiten**

Gerzymisch-Arbogast (2013:2) weist in ihrem Gutachten zu der Bezeichnung „Schriftdolmetschen“ darauf hin, dass das Schriftdolmetschen und das Simultandolmetschen viele Gemeinsamkeiten aufweisen:

- Simultaneität der Leistung
- Transferprozess (mündlicher in schriftlichen Text)
- Kohärenzherstellung und Textkondensierung angesichts der Beschränkung von Raum und Zeit
- eingeschränkte Korrigierbarkeit
- Vermeidung thematischer Sprünge in der Informationsgliederung
- Textkondensierung im Rahmen von gesprochener Sprache
- Lexikalische und syntaktische Einfachheit der gesprochenen Sprache

Auch Romero-Fresco (2011:45ff.) stellt folgende Gemeinsamkeiten fest:

- Gleichzeitiges Zuhören und Reden
- Zeitliche Beschränkung der Korrigierbarkeit
- adäquate Stimmführung

Demnach bedienen sich Simultandolmetscher\*innen und Respeaker\*innen ähnlicher Techniken. Es handelt sich bei beiden Tätigkeiten um die Übertragung eines nicht wiederholbaren, nicht fixierten, nur einmalig rezipierbaren, kaum korrigierbaren, mündlich dargebotenen Ausgangstextes (vgl. Kade 1968:35).

Aufgrund der Gemeinsamkeiten gibt es ähnliche Arbeitsabläufe. In beiden Aufgaben sollten sich die Dolmetscher\*innen im Vorhinein Gedanken machen, welches Fachvokabular oder welche Eigennamen im Ausgangstext vorkommen könnten. Diese sollten in einem Glossar beim Simultandolmetschen festgehalten bzw. dem Spracherkennungsprofil beim Respeaking beigebracht werden. Durch die Vorbereitung kann die Qualität des Zieltextes deutlich gesteigert werden (vgl. Lambourne 2006). Dieser Meinung sind auch Remael et al. (2014) und argumentieren, dass Respeaker\*innen die Qualität der Live-Untertitel verbessern, wenn ihnen das Thema vertraut ist und sie sich vorbereiten können (vgl. Remael et al. 2014:128). Käber (2016:302) sieht die Vorbereitung für die Live-Untertitelung ebenso als unerlässlich an. Neben dem Recherchieren von Namen, Fakten, Geschichten, etc. hebt er auch das Pflegen und Schönen der Stimme hervor.

Da das Respeaking – ähnlich dem Dolmetschen – kognitiv sehr fordernd ist, arbeiten Respeaker\*innen selten mehr als eine Stunde am Stück. Eine längere Dauer könnte die Qualität des produzierten Textes negativ beeinflussen. Bei zeitlich umfassenderen Einsätzen wird die Arbeit daher unter den Respeaker\*innen aufgeteilt. Teamarbeit ist deswegen auch ein wichtiger Faktor. Zudem arbeiten sowohl Respeaker\*innen als auch Dolmetscher\*innen (üblicherweise) in Dolmetschkabinen mit Mikrofon und Kopfhörern (vgl. Romero-Fresco 2019:98). Des Weiteren sind beide live tätig, was sich auch auf die Wahl der Strategie (siehe Kapitel 3.7.) auswirkt. Sie sollten den Timelag so gering wie möglich halten und auf sprachliche Richtigkeit achten (vgl. Geens 2007:105).

### **3.3.2. Unterschiede**

Im Vergleich zum Simultandolmetschen ist das Respeaking – meistens – eine intralinguale Tätigkeit. Somit gibt es keinen Transfer von einer Lautsprache in eine andere. Jedoch gewinnt das interlinguale Respeaking immer mehr an Bedeutung (vgl. Romero-Fresco/ Pöchhacker 2017) und wird zunehmend erforscht (vgl. Pöchhacker/ Remael 2019).

Im Gegensatz zum Simultandolmetschen ist der Zieltext beim Respeaking allerdings nicht für das Zielpublikum gedacht, sondern für die Spracherkennungssoftware, damit diese den Input so gut wie möglich verschriftlichen kann (vgl. Pöchhacker/ Remael 2019:135). Beim Simultandolmetschen ist der Output der Dolmetscher\*innen der Zieltext, und somit auch das

Endprodukt. Deswegen ist der Text noch immer eine Art mündliche Sprache. Beim Respeaking ist das nicht der Fall. Der gesprochene Text wird erst in einen Untertitel umgewandelt, welcher im Anschluss noch ausgebessert werden kann. Der Zieltext ist daher nicht mehr mündlich, sondern schriftlich. Dieser muss den Untertitel-Richtlinien entsprechen (vgl. Remael et al. 2014:128). Die Richtlinien können je nach Sendeanstalt unterschiedlich stark voneinander abweichen, allerdings ist man bemüht, einen Konsens zu finden (vgl. Hezel 2009).

Wie bereits erwähnt, ist die Vorbereitung sowohl beim Simultandolmetschen als auch beim Respeaking ein wichtiger Teil der Arbeit und besteht bei der Live-Untertitelung zudem darin, zu antizipieren, welche Probleme die Software beim Diktieren haben könnte. Durch das Anlegen von Makros versuchen die Respeaker\*innen diese Herausforderungen so gut wie möglich zu umgehen. Es könnten außerdem Wörter auftauchen, die vorher nicht vorbereitet werden konnten. Passende Strategien helfen dabei, das Auftreten von potenziellen Fehlern zu vermeiden (vgl. Romero-Fresco 2011:48ff.). Diese technische Komponente findet sich beim Lautsprachdolmetschen nicht.

Respeaker\*innen diktieren zusätzlich Satzzeichen und bessern die Untertitel selbst aus, sofern sie alleine arbeiten. Sie bedienen also noch zusätzlich eine Tastatur und eine Maus (vgl. Remael et al. 2014:128). Für die Tätigkeit benötigten sie Software-bezogene Fertigkeiten und Untertitel-Kenntnisse, hervorragende Grammatik- und Rechtschreibkenntnisse sowie Interpunktionssetzung. Darüber hinaus sollten sie außersprachliche Aspekte beachten, wie zum Beispiel die Farbgebung, um Sprecher\*innenwechsel sichtbar(er) zu machen. Ebenso ist die Ausrichtung bzw. Position der Untertitel zu berücksichtigen. Neben dem gleichzeitigen Zuhören und Sprechen wie beim Simultandolmetschen kommen somit noch das parallele Schreiben und Lesen beim Respeaking hinzu (Arumí Ribas/ Romero-Fresco 2008:114ff.). Romero-Fresco (2011:95) merkt jedoch an, dass auch Simultandolmetscher\*innen in Konferenzsituationen etwa PowerPoint-Präsentationen lesen oder Termini im Internet nachzuschauen haben.

Einen weiteren Unterschied sieht Romero-Fresco (2011:45ff.) in der Stimmführung. Dolmetscher\*innen sollten eine Satzmelodie wählen, die angenehm zum Zuhören ist. Respeaker\*innen sprechen dahingegen eher monoton und gleichbleibend, damit sie die Spracherkennungssoftware gut versteht. Sie achten daher mehr auf ihren Output (Intonation, Aussprache, Artikulation) und das Diktieren an sich, als dies beim Dolmetschen der Fall ist (vgl. Romero-Fresco 2019:103).

Im Kontrast zum Simultandolmetschen sollten Respeaker\*innen zudem bedenken, kurze Pausen zwischen mehreren Wörtern einzulegen, damit die Spracherkennungssoftware

den Text kontinuierlich statt in einem längeren Block anzeigt. Sie segmentieren den Ausgangstext mithilfe der Salami-Technik und diktieren in „units“, welche idealerweise fünf bis sieben Wörter lang sein sollten. Denn zu lange Sätze, die sich über mehrere Untertitel erstrecken, wären schwierig zu verstehen (vgl. Romero-Fresco 2011:108). Generell ist wichtig, dass Respeaker\*innen ihr Sprachprofil pflegen und trainieren, damit sie die Software auch gut versteht (vgl. Käber 2016).

Respeaker\*innen sollten mit überdurchschnittlicher Redegeschwindigkeit diktieren können und auf die Aussprache von kurzen Wörtern achten (vgl. Arumí Ribas/ Romero-Fresco 2008:114ff.). Angemerkt sei noch, dass sie im Gegensatz zu Simultandolmetscher\*innen nicht einfach schneller sprechen können, um den Timelag gering zu halten, falls das Tempo im Ausgangstext sehr hoch ist. Schnelleres Sprechen hätte zur Folge, dass darunter die Genauigkeit der Spracherkennung leidet und die Untertitel fehleranfälliger werden (vgl. Romero-Fresco 2019:101).

Die Einsätze der beiden Tätigkeiten weichen dahingehend ab, dass Simultandolmetscher\*innen hauptsächlich auf Konferenzen, Sitzungen, Gipfeltreffen und dergleichen arbeiten, wohingegen Respeaker\*innen Live-Sendungen im Fernsehen wie Nachrichten, Parlamentssitzungen, Sportereignisse oder Konzerte untertiteln. Aus diesem Grund unterscheiden sich auch das Zielpublikum (vgl. Marsh 2004, zit. nach Romero-Fresco 2011:46) und der Ort der Ausübung. Da das Respeaking für die Live-Untertitelung im Gegensatz zum Simultandolmetschen nicht vor Ort stattfindet, haben Respeaker\*innen keinen direkten Kontakt mit dem Zielpublikum. Deswegen obliegt es den Respeaker\*innen selbst zu kontrollieren, ob die Untertitel tatsächlich auf den Fernsehgeräten ankommen (vgl. Käber 2016:303), weil die Rückmeldemöglichkeit vom Zielpublikum fehlt (vgl. Arumí Ribas/ Romero-Fresco 2008:114ff.). Die benötigten Einstellungen müssen die Respeaker\*innen vor Sendungsbeginn auswählen und danach auf dem Rückbild überprüfen, ob die Untertitel gesendet werden (vgl. Lindner 2016:312).

### **3.3.3. Fazit der Gegenüberstellung**

Wie aus dem Vergleich ersichtlich, gibt es Gemeinsamkeiten, aber auch Unterschiede zwischen Simultandolmetschen und Respeaking. Eugeni (2010) fasst es so zusammen:

In fact, respeaking is a sort of simultaneous interpretation in the same language, which requires strong linguistic skills as well as the ability to understand and analyze complex issues and ideas and to reformulate them in real time. (Eugeni 2010:13)

Die Herausforderungen sind in beiden Tätigkeiten demnach sehr ähnlich. Es geht darum, dem Ausgangstext folgen zu können und ihn so wiederzugeben, dass er für das Zielpublikum nachvollziehbar ist (vgl. Lambourne 2006).

Beim Respeaking kommen noch technische Aspekte hinzu. Eugeni (2010:13) ist darüber hinaus der Meinung, dass weitere Kenntnisse und Fertigkeiten aus dem Bereich der SDH-Untertitel für das Respeaking benötigt werden. Dies postulieren auch Remael et al. (2014:128). Außerdem können sich die Respeaker\*innen vermehrt visueller Informationen bedienen, was dabei helfen kann, den Ausgangstext besser zu verstehen. Inwiefern sie aber aufgrund des Multitaskings beim Respeakingprozess von diesen zusätzlichen Informationen Gebrauch machen, ist unerforscht (vgl. Remael et al. 2014:128f.).

Es besteht zwar weitgehend Einigkeit darüber, dass das Respeaking eine Kombination des Dolmetsch-Prozesses und des Untertiteln ist, es wird aber als eine neue Art der audiovisuellen Translation angesehen (vgl. Romero-Fresco 2019:102).

Der vorliegende Vergleich stellt keine vollständige Aufzählung der Gemeinsamkeiten und Unterschiede von Simultandolmetschen und Respeaking dar. Vielmehr soll damit aufgezeigt werden, dass sich die Berufe stark ähneln, aber es in gewissen Bereichen deutliche Abweichungen gibt. Aufgrund der Überlappungen gibt es in der Ausbildung, welche im nächsten Unterkapitel behandelt wird, ähnliche Herangehensweisen.

### **3.4. Ausbildung und Kompetenzen**

Den Beruf Respeaking findet man in Europa zwar schon seit 2001, allerdings gibt es erst ab 2007 die erste Ausbildung im akademischen Rahmen. Die ersten Respeaker\*innen kamen von fachfremden Ausbildungen. In den letzten Jahren ist das Angebot an Untertitelungskursen in Europa gestiegen. Trotzdem stellen Respeaking-Kurse auf Hochschulebene eher noch die Ausnahme dar. Ein Großteil der Ausbildung beschäftigt sich mit den Fertigkeiten des Respeakings und der Anwendung der Spracherkennungssoftware. Ein weiterer Bereich befasst sich mit dem Erstellen von Untertiteln und den dazugehörigen Richtlinien. Damit erlangen die Teilnehmer\*innen Dolmetsch-Fähigkeiten mit dem Fokus auf dem Multitasking: simultanes Zuhören, Verstehen und Zusammenfassen des Ausgangstextes sowie Paraphrasieren und Erstellen eines Zieltextes. Romero-Fresco sieht das Multitasking beim Respeaking sogar als komplexer als beim Lautsprachdolmetschen an. Simultandolmetscher\*innen müssen laut ihm sprachlich gewandt sein und die Botschaft einer Aussage schnell verstehen. Sie hören dem Ausgangstext zu und erstellen gleichzeitig den Zieltext, während sie auch auf ihren eigenen Output achten.

Respeaker\*innen haben dieselben Aufgaben zu erledigen. Zudem müssen sie den verschriftlichten Ausgangstext kontrollieren und zusätzlich mit der Computertastatur arbeiten, um etwa Fehler zu korrigieren oder die Untertitel in die richtige Position zu bringen. Obwohl dies alles fast zur selben Zeit passiert, überlappen sich diese Aufgaben nicht komplett, was die Tätigkeit noch komplexer macht (vgl. Romero-Fresco 2019:101; 2011).

Die Frage ist, ob Simultandolmetscher\*innen auch gleichzeitig die besseren Respeaker\*innen sind. Erste Studien zu diesem Thema (vgl. Szarkowska et al. 2018) weisen auf, dass sich Dolmetschstudierende beim Respeaking besser schlagen als Studierende, die noch keine Dolmetscherfahrungen haben. Allerdings finden es Dolmetschstudierende anspruchsvoll, Satzzeichen zu diktieren und auf ihren Output und ihre Intonation zu achten. Studierende ohne Dolmetscherfahrungen haben zudem Schwierigkeiten mit der gleichzeitigen Ausübung der verschiedenen Tätigkeiten (vgl. Romero-Fresco 2019:103). Studien der Universität Warschau zeigen, dass Personen mit einer hohen Arbeitsgedächtniskapazität beim Respeaking bessere Leistungen erzielen – unabhängig davon, ob sie Dolmetscher\*innen sind oder nicht (vgl. Szarkowska et al. 2018).

Allerdings vertreten Szarkowska et al. (2018) die Meinung, dass Dolmetscherfahrung für das Respeaking von Vorteil sein kann, besonders wenn es um das kurzzeitige Abspeichern des Ausgangstextes sowie die besseren Strategien bei mehreren Informationsquellen und die präzise Weitergabe eines komplexen Ausgangstextes an die Zielgruppe geht. Chmiel et al. (2018) betonen außerdem, dass Dolmetscher\*innen im Vergleich zu Übersetzer\*innen und bilingual aufgewachsenen Menschen geübter im Umgang mit dem Auslassen von Redundanzen sind. Respeaker\*innen tun sich somit beim Produzieren eines präziseren Zieltextes einfacher. Bei der Qualität des Paraphrasierens konnten die Forscher\*innen jedoch entgegen ihrer Annahme keine Unterschiede in der Vergleichsgruppe feststellen. Romero-Fresco (2019:103) weist darauf hin, dass weitere Forschung notwendig ist, um die Ausbildung zielgerichtet zu gestalten.

Aufgrund der Gemeinsamkeiten des Simultandolmetschens und Respeakings kann davon ausgegangen werden, dass die Problemfaktoren in beiden Tätigkeiten ähnlich sind. Dafür wird im nächsten Unterkapitel zunächst auf die problemauslösenden Faktoren (*problem triggers*) nach Gile (2009) für das Simultandolmetschen eingegangen, welche dann auf das Respeaking ausgeweitet werden.

### 3.5. Effort Model nach Gile (2009) beim Respeaking und problemauslösende Faktoren

Das Effort Model von Daniel Gile (2009:157-190) beschreibt, welche Leistungen beim Simultandolmetschen gleichzeitig gefordert werden:

- Zuhören und Analysieren („The Listening and Analysis Effort - L“): Diese Dolmetsch-Teilleistung bezieht sich auf das Zuhören und Verstehen des Ausgangstextes.
- Produktion („The Production Effort - P“): Dieser Effort betrifft beim Simultandolmetschen die Zieltextproduktion. Sie beinhaltet zudem Selbstkorrektur und -überwachung.
- (Kurzzeit-)Gedächtnis („The Memory Effort - M“): Diese Teilleistung betrifft Gedächtnisleistungen und beginnt, sobald ein sprachliches Segment gehört wird, und dauert bis zur Zieltextproduktion oder bis es aus dem Gedächtnis verschwindet. Gile fügt hinzu, dass es sich hierbei nur um eine grobe Anlehnung an das Arbeitsgedächtnis („working memory“) aus der Kognitionspsychologie handelt.
- Beim Simultandolmetschen ergänzt Gile sein Modell noch um einen weiteren Effort, nämlich jenen der Koordinierung („Coordination Effort - C“): Diese Teilleistung ist für die Abstimmung der anderen drei Teilleistungen da.

Auf Basis dessen setzt sich das Simultandolmetschen laut Gile folgendermaßen zusammen:  $SI = L + P + M + C$ . Er fügt jedoch hinzu, dass die Zeichen = und + nicht im mathematischen Sinne verstanden werden sollten, sondern dass es sich hierbei um ein Zusammenspiel aller Komponenten handelt. Das Effort Model hat zum Ziel, Schwierigkeiten beim Dolmetschen aufzuzeigen. Jeder der beschriebenen Efforts benötigt Ressourcen und Entscheidungsfähigkeiten der Dolmetscher\*innen. Falls die erforderten Gesamtkapazitäten die eigentlichen, verfügbaren Kapazitäten übersteigen bzw. für einen oder mehrere Komponenten nicht ausreichend Kapazität zu Verfügung steht, weil an anderer Stelle zu viele Ressourcen aufgebraucht werden, kann die Qualität des Zieltextes darunter leiden. Da Dolmetscher\*innen oft an der Belastungsgrenze arbeiten, treten solche Fehler häufig auf. Diese Überlastung der Efforts ist auf verschiedene Problemauslöser zurückzuführen. Diese bezeichnet Gile als *problem triggers*.

Ziegler und Eichmeyer (Ziegler 2017<sup>5</sup>, zit. nach Witzel 2020:313f.) passen das Effort Model nach Gile an das Schriftdolmetschen an und fügen neben den vier genannten Efforts

---

<sup>5</sup> Ziegler, Klaus (2017). *Skript Schriftdolmetscherausbildung 2017/18. Grundlagen des Dolmetschens / Dolmetschstrategien*. SDI München/BFW Würzburg (laut Witzel 2020:323 ist das Werk unveröffentlicht).

noch zwei weitere hinzu: Outputkontrolle und Korrektur. Die Adaptierung kann als sinnvoll betrachten werden, weil diese Aspekte beim Lautsprachdolmetschen nicht gegeben sind bzw. die Outputkontrolle eine andere ist. Beim Schriftdolmetschen muss der produzierte Text nämlich zusätzlich gelesen und eventuell korrigiert werden.

Auch Martínez Pérez (2012:136) adaptiert das Effort Model nach Gile und fügt für das Respeaking ebenfalls zwei Komponenten hinzu: die technische Dimension („technical dimension“) und die visuelle Dimension („visual dimension“). Mit der technischen Dimension beschreibt er das hochtechnische Umfeld während des Respeakings. Die Respeaker\*innen müssen gleichzeitig die Spracherkennungssoftware und Untertitelungssoftware bedienen. Das heißt, sie müssen sich auf die speziellen Anforderungen beim Respeaking-Prozess (etwa Interpunktion und das monotone Sprechen beim Diktieren) konzentrieren und nonverbale und paralinguistische Äußerungen miteinbeziehen. Mit der visuellen Dimension meint er die Verarbeitung der Fernsehbilder und den erzeugten Untertiteltext, der gegebenenfalls korrigiert werden muss. Er ist der Meinung, dass diese zusätzlichen Efforts den Prozess des Zuhörens, der Analyse, der Produktion und der Gedächtnisspeicherung noch komplexer machen.

Wie bereits erwähnt, kann es bei der gleichzeitigen Ausführung verschiedener Teilleistungen zu Überlastung einzelner Bereiche und zu Fehlern kommen. Gile (2009:192) unterscheidet bei der kognitiven Überlastung zwischen chronischen Gründen („chronic reasons“) und gelegentlichen Gründen („occasional reasons“). Die chronischen Gründe beziehen sich auf die Situationen, in denen die kognitiven Kompetenzen und das deklarative Wissen der Dolmetscher\*innen für das Dolmetschen noch nicht ausreichend sind. Als gelegentliche Gründe führt Gile Probleme an, die unter bestimmten Umständen auftreten und möglicherweise zu einer kognitiven Überlastung führen. Diese können mit objektiven Faktoren verknüpft sein, die im Ausgangstext anzutreffen sind (*problem triggers*), oder mit dem Kommunikationsumfeld (erhöhter Stress, laute Umgebung, fehlendes Fachhintergrundwissen) zu tun haben. Ebenso subjektive Gründe wie kurzzeitige Aufmerksamkeitsstörung zählen darunter. Gile (2009:192f.) klassifiziert die *problem triggers* wie folgt:

1. Probleme, die erhöhte Verarbeitungskapazität erfordern:
  - a. schnelles Sprechtempo, hohe Informationsdichte des Ausgangstextes, Zahlen, abgelesene Reden
  - b. externe Faktoren wie schlechte Tonqualität, starke Akzente, falsch angewandte Grammatik, besondere lexikalische Verwendung, ungewöhnlicher Sprach- und Argumentationsstil
  - c. Eigennamen aus mehreren Wörtern, Fachtermini



- d. syntaktische Unterschiede zwischen Ausgangssprache und Zielsprache
- e. geringe Antizipierbarkeit des Ausgangstextes

2. Probleme im Zusammenhang mit mangelnder Redundanz des Inputs: kurze, wenig redundante Texte, in denen Vokale und Silben sehr ähnlich klingen können, wie etwa bei Zahlen, kurzen Namen und Akronymen

*Problem triggers* führen nicht zwangsläufig zu Problemen oder Fehlern, sondern benötigen lediglich eine erhöhte Kapazität beim Dolmetschen. Bewältigungsstrategien (*coping tactics*) können dabei Abhilfe schaffen, den Schaden in Grenzen zu halten oder zu verhindern. Gile (2009:200-211) gibt dafür folgende Beispiele: Strategien zur Verstehenssicherung (z.B. durch längeres Abwarten, Inferenzieren), Reduzierung der Belastung (z.B. durch Notizen, Segmentierung) oder Reformulierung bei der Produktion des Zieltextes (z.B. durch Paraphrasieren, Generalisierung).

Da die Auslegung der *problem triggers* sehr weit gefasst ist, ist es schwierig, eine vollständige Liste mit allen möglichen Problemauslösern aufzustellen. Mankauskiené (2018:15ff.) versucht daher in ihrer Dissertation, auf Basis eines Kommunikationsmodells von Ingram (1974/2015) und Kirchhoff (1976/2002), eine angepasste Klassifizierung der *problem triggers* vorzunehmen. Diese soll es Forscher\*innen ermöglichen, einen umfassenderen, aber dennoch gezielten Blick auf gewisse Schwierigkeiten beim Dolmetschen zu werfen. In diesem Modell werden die Problemfaktoren nach ihrer Ursache in folgende Gruppen unterteilt:

- Sender\*innenbezogene Problemauslöser (Sprecher\*innen-Akzent, Nicht-Erstsprache, schnelles Sprechen, monotone Intonation, abgelesene Texte, usw.)
- Ausgangstextbezogene Problemauslöser: Da sich alle Sprachen aus drei Systemen (lexikalisch, syntaktisch, semiologisch) zusammensetzen (Gleason 1965:91-113), können Problemauslöser, die sich auf den Ausgangstext beziehen, laut Mankauskiené (2018:17) folglich unterteilt werden:
  - Lexikalisch (Eigennamen, Zahlen, Abkürzungen, Fachterminologie)
  - Syntaktisch:
    - Phrasen (idiomatische Redewendungen, Kollokationen)
    - Sätze (syntaktische Unterschiede zwischen Ausgangs- und Zielsprache, lexikalische Dichte, lange Sätze, viele Nebensätze, Aufzählungen)
  - Semiologisch (Metaphern, Humor, Sarkasmus)
- Dolmetscher\*innen-bezogene Problemauslöser (Erfahrung, Hintergrundwissen, Kommunikationskompetenz, Ermüdung, usw.)

- Technikbezogene Problemauslöser (funktionsuntüchtige Dolmetschanlage, externe Geräusche, nicht sichtbare Redner\*innen, usw.)

Aufgrund der Unterteilung der Problemfaktoren nach der Ursache scheint das Modell von Man-kauskiené (2018) auch geeignet zu sein, um verschiedene Problemauslöser beim Respeaking ausfindig zu machen, die fernab des Lautsprachdolmetschens auftreten, wie zum Beispiel Spracherkennungsfehler oder technische Probleme mit den verwendeten Programmen.

Bevor auf Bewältigungsstrategien beim Respeaking eingegangen wird, werden im folgenden Unterkapitel noch problemauslösende Faktoren vorgestellt.

### **3.6. Problemauslösende Faktoren beim Respeaking**

Nachfolgend werden anhand von unterschiedlichen Forschungsergebnissen problemauslösende Faktoren beim Respeaking beschrieben, wengleich diese stark durch Sendungstyp und Redner\*innen beeinflusst werden. Laut Romero-Fresco (2011:123) gibt es unterschiedlich schwierige Sendungsarten zu untertiteln. Er nimmt an, dass Sportsendungen einfacher zu untertiteln sind als etwa Nachrichtensendungen oder Gesprächsrunden. Lambourne (2006) fügt hinzu, je mehr Sprecher\*innen auftreten, umso mehr könnte durcheinander gesprochen werden, und desto schwieriger würde das Respeaking werden. Diese Annahme bestätigt auch eine Studie von Szarkowska et al. (2018). Sie zeigten auf, dass Sprecher\*innen-übergreifende Dialoge am schwierigsten zu respeaken sind.

Szarkowska et al. (2019:194) fassen die problematischsten Punkte beim Respeaking zusammen und beziehen sich dabei auf Eugeni (2009) und Romero-Fresco (2011). Beim Respeaking kann es – unabhängig von der Sprachenkombination – zu folgenden Herausforderungen kommen:

- hohe Informationsdichte des Ausgangstextes
- schnelle Redegeschwindigkeit der Redner\*innen
- gleichzeitiges Sprechen verschiedener Redner\*innen
- syntaktische Fehler im Ausgangstext
- Fehler bei der Aussprache
- schlechte Qualität bzw. Störungen bei der Sendungsübertragung.

Auch Gottlieb (2002:192) schreibt, dass bei der Übertragung mündlicher Sprache in Schriftsprache die Untertitler\*innen mit folgenden Eigenheiten spontaner Rede rechnen müssen:

- Pausen, falsche Redeanfänge, Selbstkorrekturen, Unterbrechungen
- nicht beendete Sätze, grammatikalisch inakzeptable Konstruktionen

- Versprecher, Widersprüche, Ambiguitäten, Unsinn
- überlappende Rede
- dialektale oder soziolektale sprachliche Merkmale, die in der gängigen Orthografie nicht verschriftlicht werden können
- Sprecher\*innen-eigene Idiosynkrasien
- undeutliche Aussprache, sodass Wörter nicht verstanden werden können

Bisherige Studien beschäftigten sich nur marginal mit den möglicherweise auftretenden Problemen beim Respeaking. Die Forschungsarbeit von Szarkowska et al. (2019:194f.) beleuchtet sogenannte „respeaking crisis points (RCPs)“, jene Momente, in denen Respeaker\*innen mit etwaigen Problemen konfrontiert sind. Die RCPs ähneln den *problem triggers* von Gile (2009). Szarkowska et al. (2019) gehen davon aus, dass RCPs durch die kognitive Überlastung verursacht werden, welche mit der Schwierigkeit des Respeakings an sich und den Kompetenzen der Respeaker\*innen zusammenhängen könnte. Laut Szarkowska et al. (2019:195) erfordern die RCPs von den Respeaker\*innen einen größeren kognitiven Aufwand, um problematische Situationen zu umgehen. Dies könnte zu einer längeren Pause beim Output bzw. einem Leistungsabfall führen.

Szarkowska et al. (2019:195f.) untersuchen zwei Indikatoren für die kognitive Belastung beim Respeaking verschiedener Sendungsarten, worin sich auch die Inhalte, die Redegeschwindigkeit, die Sprecheranzahl und die sprachlichen Parameter unterscheiden. In der Studie konzentrieren sie sich auf die Erhebung der geistigen Belastung („mental effort“) und der psychischen Stressbelastung („psychological-stress load“) mittels Elektroenzephalogramm (EEG), um die kognitive Auslastung messen zu können. Sie gehen von der Hypothese aus, dass RCPs dann auftreten, wenn die Werte der gemessenen Indikatoren (Konzentration und/oder Frustration) deutlich vom Mittelwert abweichen. Die Forscher\*innen fügen hinzu, dass auch andere Methoden die RCPs messen könnten, wie etwa die Länge der Pausen beim Output der Respeaker\*innen oder Leistungsindikatoren (vgl. Szarkowska et al. 2019:211).

Die Forschungsergebnisse lassen den Schluss zu, dass die Abweichungen besonders wegen der Redegeschwindigkeit und schneller Redebeiträgen auftreten, welche mit den Konzentrations- und Frustrationswerten Hand in Hand gehen. Die höchsten Abweichungen bei der Frustration zeigen jene Gespräche, wo mehrere Redner\*innen gleichzeitig sprechen. Eine Ausnahme stellen Nachrichtensendungen dar, in denen zwar mehr Konzentration gefordert wird, es aber zu keinem Anstieg der Frustrationswerte kommt (vgl. Szarkowska et al. 2019:199). Dies

könnte damit zusammenhängen, dass Nachrichtensendungen meistens im semi-live Modus gesendet werden.

Im Folgenden werden die Ergebnisse der Studie für verschiedene Bereiche (politische Rede, Nachrichtensendung, Unterhaltungsprogramm und politische Diskussionsrunde) erläutert. Hierbei wird der Fokus auf das intralinguale Respeaking gelegt. Im dem Fall ist die Sprachenkombination polnisch-polnisch. Die Untersuchungen von Szarkowska et al. (2019) weisen aber auf, dass die RCPs beim interlingualen Respeaking den RCPs beim intralingualen Respeaking sehr ähneln. Neben sprachbezogenen Problemen, passende Äquivalente zu finden, gelten auch historische und geografische Bezüge sowie lange Pausen bzw. (zu) langsame Sprechgeschwindigkeit beim interlingualen Respeaking als RCPs.

Allgemein lassen sich beim intralingualen Respeaking die stärksten Abweichungen vor allem zu Beginn der Videos feststellen. Dies erklärt sich dadurch, dass die Respeaker\*innen mit einer neuen Aufgabe konfrontiert sind und sich somit auch ihre Konzentration erhöht (vgl. Szarkowska et al. 2019:199).

### **3.6.1. Politische Rede**

Bei einer intralingual gerespeakten politischen Rede ist die Konzentration am höchsten, wenn Eigennamen oder Zahlen genannt werden. Diese Informationen können nicht aus dem Kontext inferiert oder vorhergesehen werden. Zwar kann das Notieren von Zahlen beim Dolmetschen eine Hilfe sein, um den kognitiven Mehraufwand zu reduzieren, allerdings wird diese Strategie laut Szarkowska et al. (2019) beim Respeaking üblicherweise nicht verwendet. Eine andere Vorgehensweise wäre, die Zahlen so schnell wie möglich nach den Redner\*innen zu nennen, und den Satz lexikalisch umzubauen, ebenso wie die Zahlen allgemein zu halten und nur eine ungefähre Größenangabe zu machen. Das Experiment zeigt außerdem, dass es beim Respeaking aufgrund des kognitiven Mehraufwands bei der Übertragung von Zahlen zu einem Verlust der nachfolgenden Informationen kommen kann (vgl. Szarkowska et al. 2019:200).

Weitere Konzentrationshöhepunkte lassen sich feststellen, wenn Redner\*innen Anspielungen gemacht haben. Manche Respeaker\*innen geben danach zusätzliche Erklärungen bzw. Informationen, was wiederum zu einem Anstieg der Konzentration führte (vgl. Szarkowska et al. 2019:200f.).

Besonders vorgefertigte politische Reden weisen lange und komplexe Sätze auf. Die proportional hohe Informationsdichte, die langen Satzstrukturen und die kurzen Pausen können ebenso Auslöser für einen Anstieg der Konzentrationswerte sein (vgl. Szarkowska et al. 2019:201). Romero-Fresco (2011:110) rät daher, die Sätze syntaktisch zu vereinfachen und

kürzer zu gestalten. Diese Strategie ist generell vorteilhaft für das Live-Untertiteln, weil Untertitel somit kürzer und entsprechend auch einfacher zu lesen sind (vgl. Szarkowska et al. 2019:201).

### **3.6.2. Nachrichtensendung**

Nachrichtenprogramme weisen oft Merkmale wie schnelle Sprechgeschwindigkeit, hohe lexikalische Dichte, stark strukturierten Inhalt und wenige Redundanzen auf. Zudem gibt es verschiedene Sprecher\*innen und Bild und Text sind stärker miteinander verknüpft als bei anderen Formaten (vgl. Romero-Fresco 2011:126).

Eine Zunahme der Konzentrationswerte gibt es bei Nachrichtensendungen besonders dann zu finden, wenn komplexe visuelle Inhalte, wie etwa Grafiken, gezeigt werden oder wenn das Bild auf dem Bildschirm in zwei Bereiche aufgeteilt wird. Die Respeaker\*innen müssen sich wegen der verschiedenen Informationsquellen verstärkt auf den eigentlichen Sinn konzentrieren (vgl. Szarkowska et al. 2019:202f.).

Erhöhte Konzentration ist zudem erforderlich, wenn es zu Sprecher\*innenwechsel und Übertragungsänderungen kommt. Ein Anstieg lässt sich vor allem dann feststellen, wenn neben den Nachrichtensprecher\*innen noch weitere Sprecher\*innen dazukommen, die etwa von außerhalb berichten, wie zum Beispiel Reporter\*innen. Die RCPs können daher auch mit der schlechteren Tonqualität und den verschiedenen Akzenten der Interviewten zusammenhängen (vgl. Szarkowska et al. 2019:203). Aufgrund einer fehlenden vorherigen Verschriftlichung der Texte gibt es keine Möglichkeit des Ablesens, wodurch es unter anderem zu einer Verringerung der Informationsdichte, zu grammatikalischen Fehlern und einer langsameren Sprechgeschwindigkeit kommen kann.

Höhepunkte in der Konzentration lassen sich auch bei Zahlen feststellen, welche manchmal nicht nur vorgelesen, sondern auch auf dem Bildschirm eingeblendet werden. Im Normalfall kann davon ausgegangen werden, dass dies eine Erleichterung beim Respeaking ist, aber Szarkowska et al. (2019:204f.) machten hier einen weiteren RCP ausfindig. Denn teilweise entsprechen die vorgetragenen Zahlen nicht exakt jenen auf dem Bildschirm. Dies führt bei manchen Respeaker\*innen zu Verwirrung und somit zu einem Anstieg an Konzentration und Frustration.

Weiters führen atypische grammatikalische Strukturen und Einschübe zu einer Erhöhung. Aufgrund der begrenzten Zeichenanzahl und der Standzeit der Untertitel entscheiden sich die teilnehmenden Respeaker\*innen oft dazu, diese Elemente auszulassen (vgl. Szarkowska et al. 2019:205f.).

### **3.6.3. Unterhaltungsprogramm**

Beim Unterhaltungsprogramm untersuchen die Forscher\*innen jene Sendungen, in denen nicht-vorgetextete, spontane Gespräche gerespeakt werden. Die Sprechgeschwindigkeit ist langsam und die Rede beinhaltet vermehrt Merkmale mündlicher Sprache wie Zögern, falsch begonnene Sätze, Wiederholungen und Redundanzen.

Die höchste Konzentration der Respeaker\*innen lässt sich bei Eigennamen finden. Ebenso erfordern Merkmale mündlicher Sprache stärkere Aufmerksamkeit, darunter vor allem falsch begonnene Sätze und Neuformulierungen. Eine Strategie, diese Probleme zu umgehen, stellt eine kurze Décalage dar, was jedoch dazu führen könnte, unvollendete Sätze nachzusprechen. Szarkowska et al. (2019:206f.) plädieren daher dafür, den Satz erst zu respeaken, nachdem die Redner\*innen ihre Äußerung beendet haben.

Ein weiterer RCP sind Wortspiele. Paraphrasieren oder Synonyme zu wählen, hätte als Strategie bei Wortspielen nicht funktioniert, weil sonst die Anspielung verloren gegangen wäre.

### **3.6.4. Politische Gesprächsrunde**

Die Studie zeigt auf, dass politische Gesprächsrunden am schwierigsten zu respeaken sind. Dies lässt sich darauf zurückführen, dass sich die Gesprächsteilnehmer\*innen ins Wort fallen und die Sprechgeschwindigkeit sehr hoch ist. Auch Namen stellen eine Herausforderung dar. Eine gute Vorbereitung könnte dabei helfen, dieses Problem zu umgehen. Ebenso könnten Äußerungen über kontroverse Themen ein RCP darstellen (vgl. Szarkowska et al. 2019:207f.).

### **3.6.5. Zusammenfassung**

Die Studie weist auf, dass besonders die Redegeschwindigkeit mit den „respeaking crisis points“ zusammenhängt, sowohl bei sehr schneller als auch langsamer Sprechgeschwindigkeit. Weitere Faktoren sind die Anzahl der Sprecher\*innen, Sprecher\*innenwechsel und überlappende Redebeiträge. Ebenso können Zahlen, Eigennamen, komplexe syntaktische Satzstrukturen sowie Nebensätze, mehrere zeitgleiche visuelle Informationsquellen, Wortspiele und Andeutungen als RCPs bezeichnet werden (vgl. Szarkowska et al. 2019:211). Außerdem lassen sich folgende RCPs feststellen: schlechte Tonqualität, verschiedene Akzente, kontroverse Äußerungen und Neuformulierungen bzw. falsch begonnene Sätze (vgl. Szarkowska et al. 2019).

Die Autor\*innen heben hervor, dass es trotz derselben Aufgaben bei den Respeaker\*innen zu unterschiedlichen kognitiven Belastungen kommt. Mithilfe von Strategien, die durch Übung und Erfahrung erlangt werden können, kann die Belastung reduziert und die Leistung verbessert werden. RCPs führen also nicht gleich zu einem Leistungsschwund beim Respeaking

(vgl. Szarkowska et al. 2019:211f.). Auf die Strategien beim Respeaking wird im folgenden Unterkapitel eingegangen.

### **3.7. Strategien beim Respeaking bzw. Schriftdolmetschen**

Im Kapitel 3.5. wurden die Bewältigungsstrategien (*coping tactics*) nach Gile (2009) kurz vorgestellt. Witzel (2020:307-310) vergleicht auf Basis der aufgestellten Dolmetschstrategien von Gile (2009) und Kalina (1998) die Strategien für das Lautsprachdolmetschen und das Schriftdolmetschen. Laut Witzel (2020) wenden Schriftdolmetscher\*innen Verstehensstrategien, Planungsstrategien, Strategien der Zieltextproduktion und Monitoringstrategien an:

- Verstehensstrategien:
  - Vorbereitung und Vorwissen in Dolmetschsituation aktivieren (Lesen und Hören von Fachtexten, Systeme mit relevanter Fachterminologie pflegen)
  - Inferenzieren (Schlussfolgern)
  - Antizipieren (Vorwegnehmen)
- Planungsstrategien:
  - Segmentieren des Textes; Herstellen von Hierarchien und Bezügen
  - Textzusammenhänge erfassen; abgeschlossene Sinneinheiten erkennen
  - Absatzsetzung für das bessere Schriftbild beim Lesen
  - Kurze Sätze bilden
  - Décalage
- Zieltextproduktion:
  - Transkodierung (etwa bei Zahlen und Namen)
  - Suche von Äquivalenten
  - Kohärenzherstellung
  - Notstrategien wie Kompression, Verallgemeinerung, Auslassung
  - Präsentation
- Monitoringstrategien:
  - Outüberprüfung auf Kohärenz und Richtigkeit
  - simultanes Zuhören und Korrigieren des Transkripts

### **3.8. Aktuelle Untersuchungen zum Respeaking in der Live-Untertitelung**

Dass das Respeaking im Bereich der Translationswissenschaft immer häufiger einen Gegenstand von Untersuchungen darstellt, zeigt sich auch daran, dass vermehrt Masterarbeiten zu dem Thema verfasst werden. Auf zwei davon soll im Folgenden kurz eingegangen werden.

Zunächst ist hierbei die Arbeit von Jorda (2021) zu nennen, die sich mit Problemlösern und Strategien beim intralingualen Respeaking (deutsch-deutsch) für die Live-Untertitelung beschäftigt. Mittels Fallstudie untersuchte die Autorin eine politische Talkshow und eine Magazinsendung und konnte genreübergreifend folgende Problemlöser ausfindig machen: Eigennamen, Fachtermini, hohe Sprechgeschwindigkeit, schnelle Sprecher\*innenwechsel, abgebrochene und redundante Sätze sowie komplexe Satzstellungen. In der Talkshow fielen auch Zahlen und Sprecher\*innenüberlappungen darunter. Eine\*r der zwei Studienteilnehmer\*innen hatte zudem Probleme mit der Spracherkennungssoftware, was die Respeaking-Leistung beeinträchtigte. Diese Problemfaktoren führten zu Auslassungen, Sinnverschiebungen, verzögerter Untertitelproduktion und Versprechern beim Diktieren. Die genannte Untersuchung zeigte ebenso auf, dass dieselben Problemfaktoren nicht zu denselben Problemen bei den Respeaker\*innen führen, welche zur Problemlösung – je nach persönlicher Erfahrung – unterschiedliche Strategien anwenden. Hierzu zählen die Salamitechnik, das Anpassen des Timelags sowie Paraphrasieren und Tilgung von Redundanzen.

Im Rahmen der zweiten Masterarbeit (Stehle 2021) befasst sich der Autor mit der Anwendung des NER-Modells (Romero-Fresco/ Martínez Pérez 2015) zur Qualitätsbewertung von intralingualen Live-Untertiteln (deutsch-deutsch). Er konnte aufzeigen, dass das NER-Modell für die Bewertung der Genauigkeit der Untertitel, die mittels Respeaking erstellt wurden, gut geeignet ist, jedoch Schwierigkeiten im Ausgangstext, Untertitelverzögerungen und die Kennzeichnung von Sprecher\*innenwechseln nicht hinreichend berücksichtigt. Außerdem kommt es in der Live-Untertitelung zu einer Textreduktion, allerdings hat dies nicht zwingend eine Auswirkung auf die Qualität der Untertitel. Auch in dieser Studie führten technische Probleme mit der Spracherkennung zu einer reduzierten Untertitelqualität. Zudem stellte sich heraus, dass kurze Sequenzen aufgrund schneller Geschwindigkeit bzw. Sprecher\*innenwechsel häufig ausgelassen und die meisten Untertitel vor dem Senden noch editiert werden mussten, was in einer Verzögerung der gesendeten Live-Untertitel resultierte.



## 4. Fragestellungen und Methodik

In diesem Kapitel werden die Fragestellungen sowie die Methodik der Masterarbeit beschrieben. Zuerst wird auf den Zugang des Forschers zum Feld eingegangen.

### 4.1. Zugang zum Feld

Der Verfasser der Masterarbeit weist mehrjährige Berufserfahrung als Respeaker auf und hat Sendungen unterschiedlichster Art bearbeitet. Durch die professionelle Tätigkeit als Live-Untertitler besteht Zugang zum Forschungsfeld. Der Forscher ist daher selbst als Experte zu betrachten und bringt eigene Erfahrungen in die Analyse mit ein. Des Weiteren werden erfahrene Kolleg\*innen für Expert\*inneninterviews herangezogen. Expert\*innen weisen sich dadurch aus, dass sie ihr Wissen verallgemeinern und konkrete Erfahrungen liefern (vgl. Helfferich 2014:570). Dies fließt ebenso in die Analyse mit ein.

### 4.2. Aufstellung live-untertitelter Sendungen

Im ersten Abschnitt des empirischen Teils soll eine Aufstellung der Live-Sendungen vorgenommen werden, die in einer Untertitelungsfirma produziert werden. Es soll herausgefunden werden, welche Arten von Sendungen die Respeaker\*innen bearbeiten.

Zwar kann man Sendungen in wissenschaftliche, ökonomische und rechtlich-politische Bereiche klassifizieren, allerdings geht diese Unterteilung zu sehr ins Detail und deckt sich nicht mit den Einschätzungen des Publikums (vgl. Gehrau 2001:267). Vielmehr soll deswegen eine allgemeine, einfache Einteilung für die Aufstellung herangezogen werden. Die ARD/ZDF-Programmanalyse<sup>6</sup> unterscheidet zwischen vier redaktionellen Programmsparten und den dazugehörigen Sendungsformaten (vgl. Maurer et al. 2021):

1) journalistische Information: Nachrichtensendungen inklusive Wetterberichte; Magazinsendungen wie Frühstücksmagazine, Ratgeber- und Servicemagazine, Boulevardmagazine, allgemeine Magazine; Reportagen, Dokumentationen; Interview- und Talkformate; sonstige journalistische Formate wie Ansprachen und Übertragungen;

2) Sportsendungen: Übertragungen, Reportagen, Dokumentationen, Nachrichten- und Magazinsendungen;

---

<sup>6</sup> Die Programmanalyse der öffentlich-rechtlichen Fernsehanstalten ARD und ZDF hat zum Ziel, die strukturelle und inhaltliche Entwicklung der deutschen Fernsehprogramme aufzuzeigen (vgl. <https://www.ard-media.de/media-perspektiven/studien/programmanalyse>, Stand: 05.05.2022).

3) nonfiktionale Unterhaltung und Reality-TV: Show-, Quiz-, Musiksendungen, etc.; Satire, Kabarett, Comedy; Reality TV wie Daily Talk, Doku-Soap, Real-Life-Inszenierung, Casting-Show, Scripted Reality;

4) fiktionale Unterhaltung: Filme, Fernsehspiele, Serien;

Zudem zählt die Programmanalyse noch eine Restkategorie mit sonstigen Sendungen (wie religiöse Verkündigungssendungen und Gottesdienste) und nichtredaktionelle Programmelemente (wie Werbung oder Programmpromotion) hinzu (vgl. Maurer et al. 2021).

Für die Aufstellung der Live-Sendungen wird der Dienstplan einer Untertitelungsfirma vom September 2020 herangezogen und nach den oberhalb angeführten Programmsparten und Sendungsformaten eingeteilt. Da die meisten Sendungen in einem unterschiedlichen Tages- oder Wochenrhythmus oder auch nur einmalig ausgestrahlt werden, entschied sich der Forscher für einen zweiwöchigen Untersuchungszeitraum (16. bis 30. September 2020), um ein gesamtlicheres Abbild der Sendungen zu gewinnen. Folgende Fragen sollen geklärt werden:

- Wie lange dauern die Sendungen, die von Respeaker\*innen Untertitelt werden?
- Wie sieht die prozentuale Verteilung von reinen live, vorproduzierten bzw. semi-live und gemischten (live und semi-live) Untertiteln der behandelten Sendungen aus?
- Mit welcher Programmsparte sind die Respeaker\*innen am häufigsten konfrontiert?
- Mit welchem Sendungsformat sind die Respeaker\*innen am häufigsten konfrontiert?
- Welche Sendungsformate werden mit welcher Methode (live/semi-live/gemischt) Untertitelt?

Für die Analyse wurde ein Raster mit verschiedenen Kategorien entworfen (siehe Abb. 3), in dem die Sendungsdauer, gruppiert in kurz (1 bis 15 Minuten), mittellang (16 bis 59 Minuten) und lang (60+ Minuten), aufgelistet wird. Des Weiteren werden die Sendungsinhalte kurz beschrieben, um die Sendungen nach dem Schema von Maurer et al. (2021) in Programmsparte und Sendungsformat einzuteilen. Außerdem wird zwischen as-live bzw. Live-Sendungen und vorab aufgezeichneten Programmen unterschieden. Bei vorab aufgezeichneten Sendungen liegt die sendefertige Fassung weit vor Ausstrahlungsbeginn vor. Eine Ausnahme stellt hierbei der Wetterbericht dar. Zwar kommen manche dieser Videos relativ kurzfristig, diese haben aber mit nur zwei Minuten eine relativ kurze Dauer. Es bleibt im Normalfall ausreichend Zeit, das Wetter vor Sendungsbeginn komplett vorzubereiten. Zudem wird im Raster angegeben, um welche Art von Untertiteln es sich bei den Sendungen handelt, mit einer Differenzierung zwischen live-produzierten Untertiteln, vorproduzierten (also fertig Time-codierten) bzw. semi-

live (die Timecodes werden während des Sendens automatisch von der Untertitelungssoftware generiert) Untertiteln und gemischten (entspricht live und semi-live) Untertiteln. Erfahrungsgemäß kann dies bei den gleichen Sendungen unterschiedlich ausfallen, je nachdem wie viel (brauchbares) Material vor einer Sendung verfügbar ist. Für die Aufstellung wird vom Regelfall ausgegangen. Am Ende soll sich herauskristallisieren, welche Arten von Sendungen die Respeaker\*innen (besonders häufig) live untertiteln, wie sich diese prozentuell untereinander aufteilen und welcher Modus dabei zum Tragen kommt.

Minuten	Beschreibung der Sendung	Sendungsformat	Programm- sparte	Livesendung und as- live	vorab aufgezeichnete Sendung	reine live UT	vorproduzierte bzw. semi- gemischte live UT	Infos
45	unterhaltsame Ratgebersendung mit Koch- und Backrezepten, Hilfestellungen im Alltag und kreativen Ideen für zu Hause	Ratgeber- und Servicemagazin	Informationssendung	x			x	

Abbildung 3: Ausschnitt des selbst entworfenen Rasters mit den verschiedenen Kategorien

### 4.3. Workflow und Problemfaktoren bei der Live-Untertitelung

Der zweite Abschnitt des Empirieteils beschäftigt sich mit dem Workflow beim Live-Untertiteln und den damit verbundenen Problemfaktoren. Auf Basis des Dolmetsch-Prozessmodells von Kalina (2005)<sup>7</sup> entwerfen Pöchhacker/ Remael (2019:136f.) ein Prozessmodell für das interlinguale Live-Untertiteln. Die Aufgabenunterteilung beim interlingualen Live-Untertiteln unterscheiden sie folgendermaßen: Pre-Process, Peri-Process und Post-Process (siehe Abb. 4). Die einzelnen Komponenten können als Ausgangspunkt für den Workflow beim intralingualen Respeaking für die Live-Untertitelung herangezogen werden. Den Arbeitsvorgang kann man in Vorbereitung, Live-Sendung und Nachbearbeitung aufgliedern.

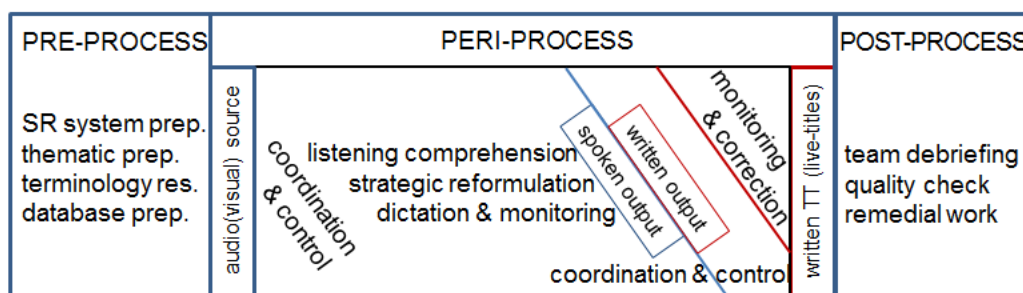


Abbildung 4: Prozessmodell beim interlingualen Live-Untertiteln (Pöchhacker/ Remael 2019:137)

<sup>7</sup> Beim literarischen Verweis dürfte den Autor\*innen Pöchhacker/ Remael (2019) ein Fehler bei der Angabe der Jahreszahl unterlaufen sein. Das beschriebene Prozessmodell ist in Kalina (2005) zu finden und nicht wie angegeben in Kalina (2000). Interpreting competences as a basis and a goal for teaching. *The Interpreters' Newsletter* 10, 3-32.

- Pre-Process: Dieser Abschnitt umfasst Dolmetschfähigkeiten und -kompetenzen, Informationsbeschaffung und -aufbereitung sowie die Koordination oder Kooperation mit Kolleg\*innen (vgl. Kalina 2005:778). Die Vorbereitung im Pre-Process ist beim interlingualen Live-Untertiteln von großer Relevanz. Dazu zählt generell die optimale Bedienung der Spracherkennungssoftware (Hardware-Einstellungen; Spracherkennungsprofil erstellen, pflegen und trainieren; Anlegen von Makros; Anpassen der jeweiligen internen Richtlinien). Des Weiteren erfolgt die thematische, sprachliche und terminologische Recherche für eine spezifische Sendung. Vorbereitete Inhalte werden in die Spracherkennungssoftware eingepflegt (vgl. Pöchhacker/ Remael 2019:136f.). Diese Teilkomponente findet beim intralingualen Respeaking im Zuge der Vorbereitungsphase für eine Sendung statt und dauert an bis zu Sendungsbeginn.
- Peri-Process: Zu diesem Abschnitt zählt der Dolmetsch- bzw. Respeakingvorgang selbst (siehe Kapitel 3), das Anwenden von Strategien sowie die Zusammenarbeit im Team (vgl. Kalina 2005:778; Pöchhacker/ Remael 2019:137). Diese Komponente umfasst somit das intralinguale Live-Untertiteln während der Sendung (alleine oder im Team) mittels der zuvor beschriebenen Respeaking-Methoden.
- Post-Process: Zu diesem Abschnitt gehören etwa das Nachbesprechen mit Teammitgliedern, um Lösungen für Probleme zu ermitteln, Qualitätskontrolle und Vorbeugungsmaßnahmen sowie Verbesserungen bzw. Trainieren der Spracherkennungssoftware für zukünftige Einsätze (vgl. Pöchhacker/ Remael 2019:137). Für den Workflow beim intralingualen Respeaking kann dies auf die Nachbearbeitung einer Sendung umgemünzt werden.

Anhand des Modells von Pöchhacker/ Remael (2019) zum interlingualen Live-Untertiteln und einer Selbstreflexion soll der Workflow beim intralingualen Respeaking einer konkreten Sendung wiedergegeben werden. Zusätzlich werden zwei Respeaker\*innen zum Ablauf der gleichen Sendung interviewt, um die Ergebnisse der Selbstreflexion mit den Erkenntnissen der Expert\*inneninterviews abzugleichen und zu ergänzen.

Überdies soll mithilfe der qualitativen Interviews herausgefunden werden, mit welchen Problemfaktoren die Respeaker\*innen beim Live-Untertiteln konfrontiert sind. Qualitative Interviews eignen sich, um authentische Informationen aus dem Bereich der Befragten zu gewinnen. Sie weisen wenig Standardisierung auf, werden mündlich-persönlich geführt, bestehen aus offenen Fragen und der Interviewstil ist neutral bis weich (vgl. Lamnek 2010:301-325). Für die

vorliegende Arbeit entschied sich der Forscher für ein teilstandardisiertes Expert\*inneninterview.

Im Vorfeld der Durchführung der Interviews wurde ein Leitfaden (siehe Anhang) entworfen. Die Fragen sind in drei große Themenbereiche (Vorbereitung, Live-Sendung, Nachbearbeitung) gegliedert, um die Workflow-Phasen getrennt voneinander zu erfassen. Sie sind möglichst offen gehalten, damit die interviewten Expert\*innen ein relativ großes Ausmaß an Freiheit für die Beantwortung haben. Abschließend haben die Interviewpartner\*innen noch die Möglichkeit, sonstige Anmerkungen zum Workflow beim Live-Untertiteln zu machen. Ziel der Befragung ist, festzuhalten, welche Arbeitsabläufe die Respeaker\*innen beim intralingualen Live-Untertiteln durchgehen und mit welchen Problemfaktoren sie in den Phasen konfrontiert sind.

Vor den Einzelinterviews wurden die Expert\*innen über das Thema der Masterarbeit informiert und es wurde die Zustimmungserklärung zur Aufnahme und Datenschutzmitteilung eingeholt. Die Interviews fanden am Arbeitsplatz der Live-Untertitler\*innen statt und dauerten rund 15 bis 20 Minuten. Teilweise wurden gewisse Fragen der Situation angepasst und hinzugefügt bzw. absichtlich ausgelassen, weil sie im Zuge des Gesprächs schon beantwortet worden waren. Anschließend erfolgte eine Transkription der Interviews (siehe Anhang), worin die Interviewpartner\*innen in IP1 und IP2 umbenannt wurden, um deren Anonymität zu wahren. Die Transkriptionen wurden geglättet, und die gesprochene Sprache der Schriftsprache angepasst. Nonverbale Elemente spielten für die Auswertung keine Rolle und fanden daher keine Berücksichtigung. Im folgenden Kapitel findet eine Darstellung der gewonnenen Ergebnisse statt.

## 5. Ergebnisse

In diesem Kapitel werden die Ergebnisse der Masterarbeit präsentiert. Im ersten Abschnitt wird zunächst auf die Sendungskategorien eingegangen und beschrieben, welche Sendungen live untertitelt werden. Im zweiten Teil wird erläutert, wie eine ausgewählte Sendung live untertitelt wird. Außerdem werden die Problemfaktoren in den verschiedenen Workflow-Phasen aufgezeigt.

### 5.1. Aufstellung live-untertitelter Sendungen

Im ausgewählten Zeitraum (16. bis 30. September 2020) standen 275 Sendungen auf dem Dienstplan, die live untertitelt werden sollten. Insgesamt entspricht dies einer Minutenanzahl von 5706. Werden die Sendungen in kurze (1 bis 15 Minuten), mittellange (16 bis 59 Minuten) und längere (60+ Minuten) Sendungen unterteilt, zeigt sich (siehe Abb. 5), dass kurze Sendungen (167 von 275) mit 61 Prozent<sup>8</sup> den Großteil aller live-untertitelten Sendungen ausmachten. Mittellange Sendungen (103 von 275) hatten einen Anteil von 37 Prozent, wohingegen Sendungen mit über einer Stunde Sendezeit (5 von 275) nur zwei Prozentpunkte ausmachten.

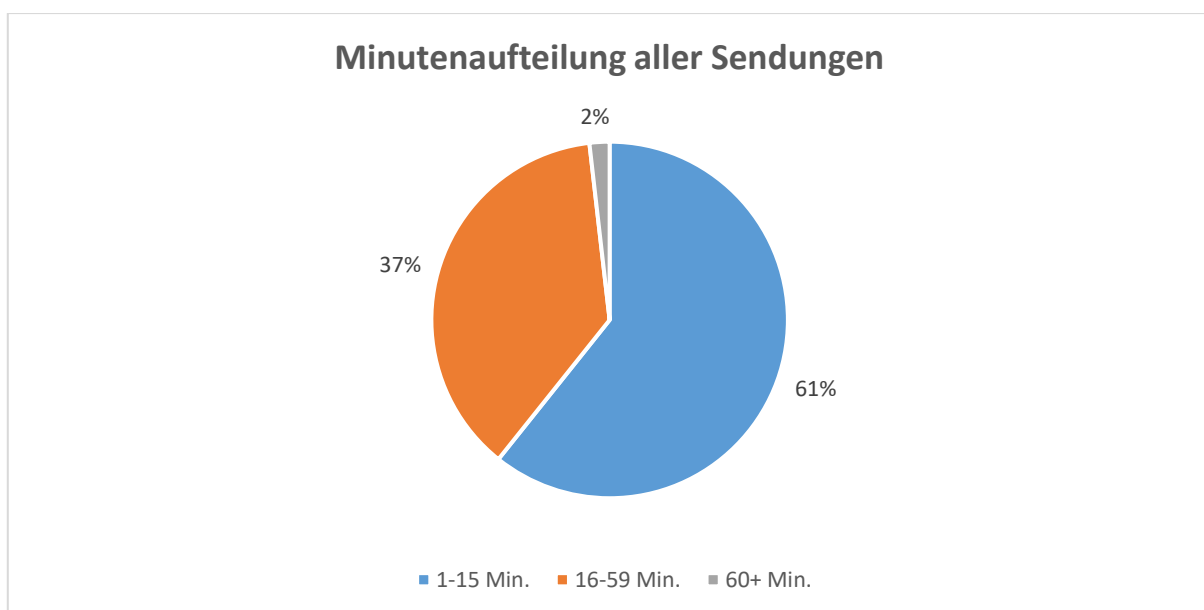


Abbildung 5: Minutenaufteilung aller Sendungen in Prozent

Von den 275 Sendungen wurden 234 as-live bzw. live ausgestrahlt. Dies entspricht 85 Prozent aller Sendungen. Die restlichen Sendungen (15 Prozent) wurden vorab aufgezeichnet. Diese mussten aufgrund der erst kurzfristig fertig gewordenen Schnittfassung trotzdem live untertitelt

<sup>8</sup> Die angegebenen Prozentpunkte werden auf ganze Zahlen gerundet.

werden. Man könnte von der Annahme ausgehen, dass vorab aufgezeichnete Sendungen zumindest im semi-live Modus Untertitelt werden können, weil das sendefertige Video vor Sendebeginn vorliegt. Diese Vermutung bestätigt sich auch für den untersuchten Zeitraum. Von den 41 vorab aufgezeichneten Sendungen wurden 38 mit vorproduzierten bzw. semi-live Untertiteln gesendet. Drei Sendungen mussten hingegen komplett live Untertitelt werden, weil weder Texte noch MAZen<sup>9</sup> (mit ausreichend Bearbeitungszeitraum) zur Verfügung gestellt werden konnten. In 93 Prozent der Fälle gab es bei vorab aufgezeichneten Sendungen somit Vorbereitungs-material. Es zeigt sich, dass die Sendeanstalten bemüht sind, vorhandenes Material bereit zu stellen, was wiederum hilft, die Herausforderungen beim Live-Untertiteln zu reduzieren. Dadurch kann auch für eine bessere Untertitel-Qualität gesorgt werden. Dass Respeaker\*innen mit Vorbereitungs-material rechnen können, lässt sich ebenso feststellen, wenn man alle Sendungen mit Live-Untertiteln in reine live-produzierte, vorproduzierte bzw. semi-live und gemischte Untertitel unterteilt: Die Untertitel von 189 Sendungen (69 Prozent) wurden im semi-live Modus gesendet, 78 (28 Prozent) waren gemischt und nur acht (drei Prozent) komplett live (siehe Abb. 6). Die überwiegende Mehrheit aller live-untertitelten Sendungen hatte somit zumindest Anteile vorbereiteter Untertitel. Das heißt, dass die eigentlichen Respeaking-Teile kürzer ausfielen.

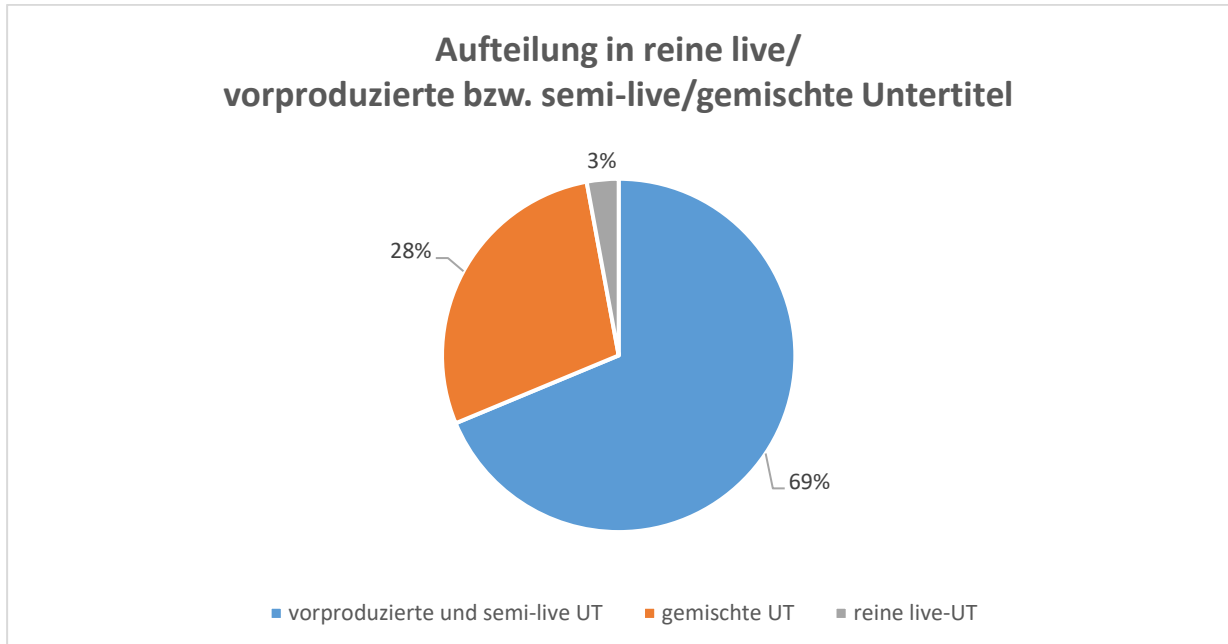


Abbildung 6: Aufteilung der Sendungen in reine live/vorproduzierte und semi-live/gemischte Untertitel in Prozent

<sup>9</sup> MAZ steht für „Magnettaufzeichnung“. Dabei handelt es sich um fertig aufgezeichnete Einspielvideos, für die die Untertitel (meist) schon vorbereitet werden können und diese nur noch mit Enter gesendet werden müssen.

Wenn man die Sendungen nach Programmsparte kategorisiert (Unterteilung in Informationssendung, Sportsendung, nonfiktionale Unterhaltung, fiktionale Unterhaltung), lässt sich ablesen, dass von den 275 untersuchten Sendungen fast ausschließlich Informationssendungen (95 Prozent bzw. 262 Sendungen) live untertitelt wurden (siehe Abb. 7).



Abbildung 7: Aufteilung aller Sendungen nach Programmsparte in Prozent

Sportsendungen machten vier Prozent (zehn Sendungen) aus. Dem folgen Sendungen mit nonfiktionaler Unterhaltung und einem Prozentpunkt (drei Sendungen). Fiktionale Unterhaltung wie Filme oder Serien finden sich in der Grafik nicht. Dies ist damit zu erklären, dass in solchen Fällen die sendefertige Fassung schon lange vor Sendungsausstrahlung vorliegt und demnach die Untertitel im Voraus vorproduziert werden können. Fiktionale Unterhaltung spielt für die Live-Untertitelung daher keine Rolle.

Sieht man sich die Minutenaufteilung (siehe Abb. 8) nach Programmsparte an, stellt sich das Bild etwas anders dar: Informationssendungen machen zwar mit 5066 Minuten (89 Prozent) bei der Minutenaufteilung aller Sendungen nach Programmsparte noch immer den überwiegenden Teil aus, allerdings gewinnen Sportsendungen und nonfiktionale Unterhaltungssendungen (je 320 Minuten bzw. sechs Prozent) etwas an Prozentpunkten dazu.



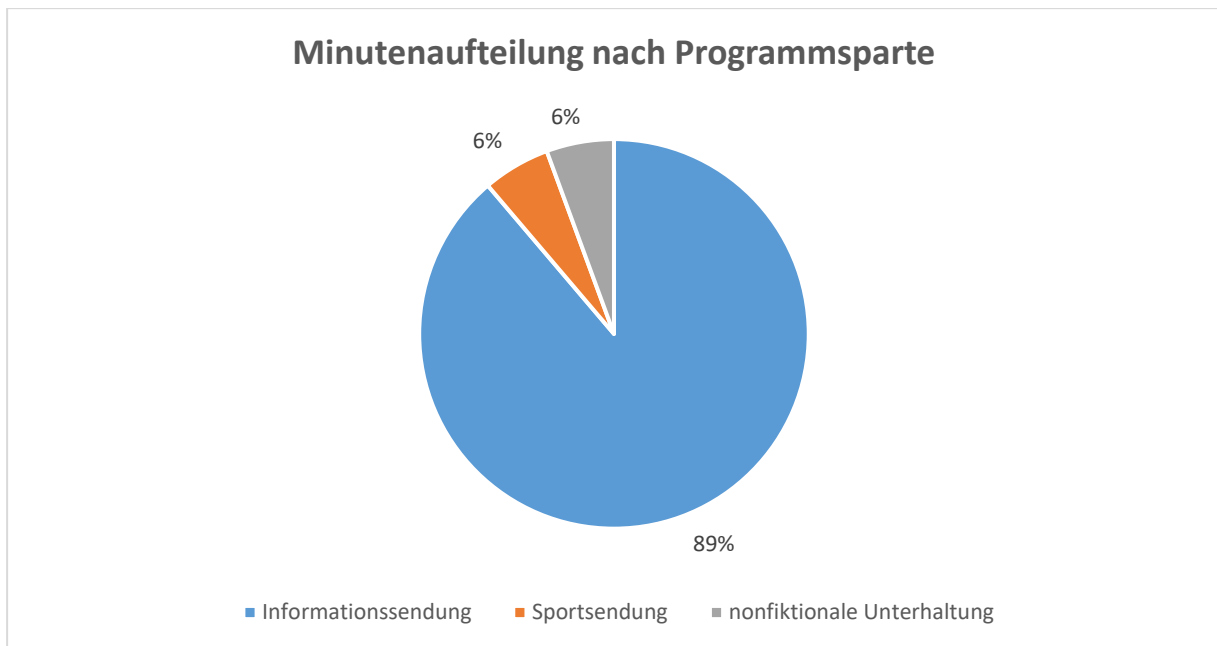


Abbildung 8: Minutenaufteilung nach Programmsparte in Prozent

Grund dafür ist, dass Informationssendungen meist aus Nachrichtensendungen bestehen und diese im Vergleich zu Sportsendungen oder Unterhaltungssendungen nur eine relativ kurze Sendungslänge aufweisen. Rechnet man sich den Durchschnitt der Sendeminuten nach Programmsparte aus, kann man feststellen, dass Informationssendungen mit 19 Minuten pro Sendung den kürzesten Anteil haben. Etwas länger dauern Sportsendungen mit 32 Minuten pro Sendung. Nonfiktionale Unterhaltungssendungen heben sich mit einer Länge von 107 Minuten pro Sendung von den anderen Programmsparten deutlich ab.

Die Programmsparten sollen einen groben Überblick über den Inhalt der live-untertitelten Sendungen bieten. Die folgende Unterteilung in verschiedene Sendungsformate soll nun einen detaillierteren Einblick in die thematischen Unterschiede live-untertitelter Sendungen geben. Die 275 untersuchten Sendungen können in folgende Formate unterteilt werden:

- Eine Dokumentation mit anschließendem Talk: Bei der Sendung ging es um eine spontan gesendete Dokumentation zum Thema Rechtsextremismus. Anschließend gab es eine Gesprächsrunde mit Fachexpert\*innen, Politiker\*innen und weiteren Gästen.
- 77 Magazinsendungen: Hierbei handelte es sich um verschiedene Arten von Magazinsendungen wie Informations- und Nachrichtenmagazine, Ratgeber- und Servicemagazine und Wissensmagazine. In diesen Sendungen ist ein dementsprechend breites Themenspektrum zu finden (Wirtschaft, Politik, Kultur, Gesellschaft, Haushalt, Garten, Finanzen, Handwerk, Computer, Ernährung, Kulinarik, Recht, Gesundheit, Sport, etc.).
- Zwei Musiksendungen: Hierbei dreht sich die Sendung um eine neue Art von Gesangsshow, bei der Kandidat\*innen anfangs nur zu sehen und nicht zu hören sind.

- 194 Nachrichtensendungen inklusive Wetterberichte: Diese Sendungen umfassen Regionalnachrichten über aktuelle Ereignisse sowie das Regionalwetter.
- Eine Quizsendung: In der Sendung mussten die Teilnehmerinnen verschiedene Aufgaben lösen und am Ende wurde eine Gewinnerin gekürt.

Wie in Abbildung 9 ersichtlich, machten Nachrichten und Wetter mit 70,5 Prozent den größten Anteil aller live-untertitelten Sendungen aus. Dem folgten Magazinsendungen mit 28 Prozent. Dokumentationen (0,4 Prozent), Musiksendungen (0,7 Prozent) und Quizsendungen (0,4 Prozent) hatten nur einen verschwindend geringen Anteil aller Sendungen.

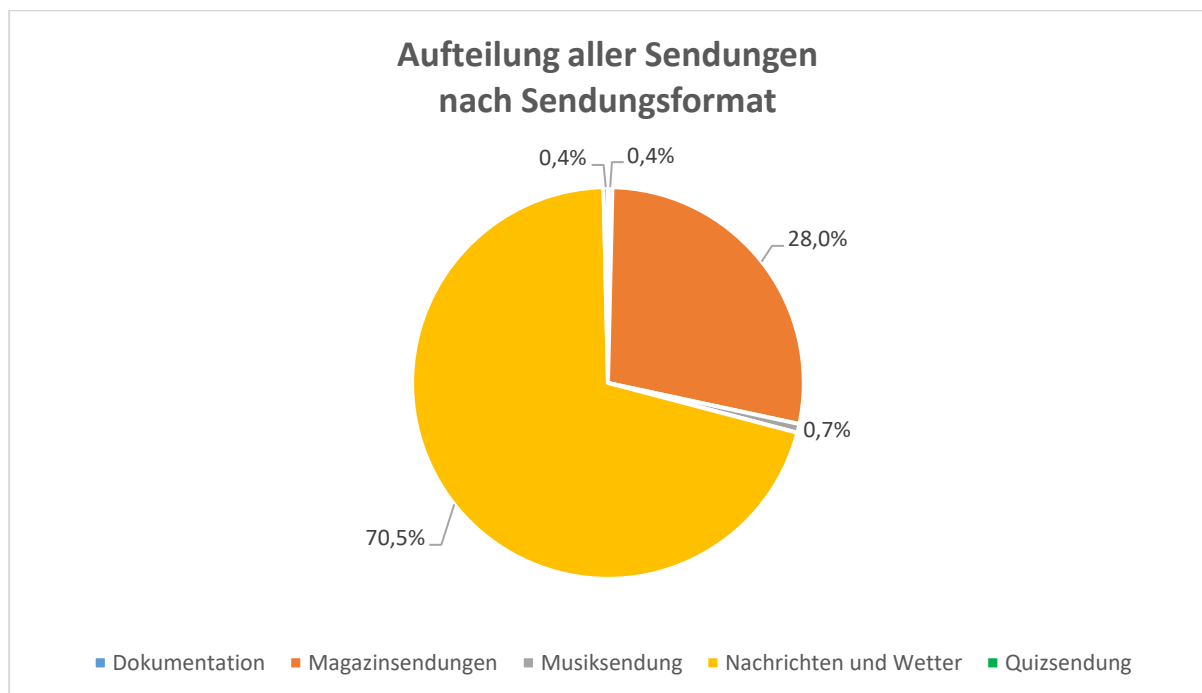


Abbildung 9: Aufteilung aller Sendungen nach Sendungsformat in Prozent

Allerdings ist festzustellen, dass sich die prozentuale Verteilung der Sendungsformate bei der Minutenanzahl deutlich verändert. Bei der Sendezeit überholen die Magazinsendungen (3442 Minuten) die Nachrichtensendungen inklusive Wetterberichte (1824 Minuten). Auch Musiksendungen (200 Minuten), Dokumentationen (120 Minuten) und Quizsendungen (120 Minuten) gewinnen leicht an Prozentpunkten hinzu (siehe Abb. 10).

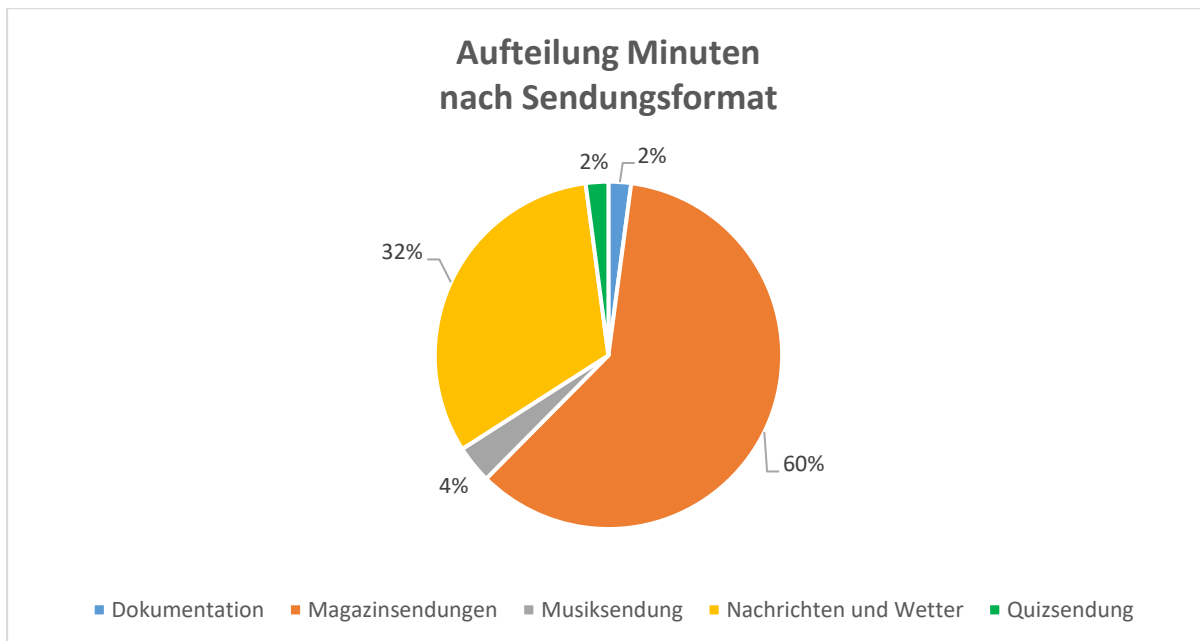


Abbildung 10: Aufteilung der Minuten nach Sendungsformat in Prozent

Teilt man die Sendungsformate in vorproduzierte bzw. semi-live Untertitel, reine live-gesendete Untertitel und gemischte Untertitel auf, ergibt sich folgendes Bild: Bei 78 Sendungen wurden gemischte Untertitel gesendet. Den größten Anteil dabei machen Magazinsendungen (65 Sendungen) mit rund 83 Prozent aus. Dem folgen Nachrichten und Wetter (zwölf Sendungen) mit rund 15 Prozent und eine Quizsendung (rund ein Prozent) (siehe Abb. 11). Die Musiksendungen und die Dokumentationsendung weisen keine vorbereiteten Untertitel auf.

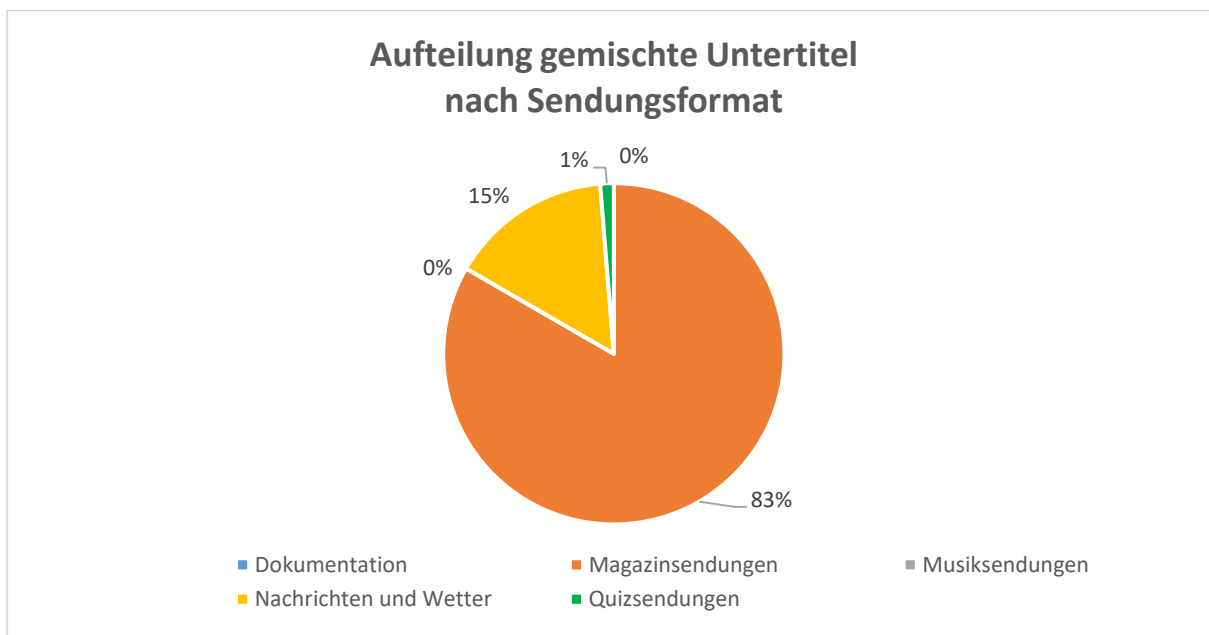


Abbildung 11: Aufteilung gemischte Untertitel nach Sendungsformat in Prozent

Diese Aufteilung ist damit zu begründen, dass Magazinsendungen oft mehrere Live-Teile haben, bei denen Gäste zu Wort kommen. Diese Abschnitte müssen mittels Respeaking untertitelt

werden. Videos und An- bzw. Abmoderationen lassen sich dahingegen vor Sendungsbeginn vorbereiten und im semi-live Modus abschicken. Teilweise kommt es aber auch vor, dass Interviews kurz vor Beginn der Sendung voraufgezeichnet werden. Sie sind dann in etwa 30 bis 60 Minuten vor Sendungsbeginn abrufbar. Hierbei hängt es davon ab, ob die Respeaker\*innen genügend Zeit haben, das Gespräch vor Sendungsbeginn zu untertiteln, und die Untertitel im Idealfall im semi-live Modus abschicken können. Da dies selbst bei der gleichen Sendung immer unterschiedlich sein kann, sind die Untertitel dieser Magazinsendungen als gemischte Untertitel zu kategorisieren. Bei gewissen Nachrichtensendungen gibt es ebenfalls Gäste, die live zu Wort kommen. Beitragstexte für Videos sind jedoch als Vorbereitungsmaterial verfügbar. Dies ist auch bei manchen Wetterberichten der Fall. Die MAZen lassen sich mit dem zur Verfügung stehenden Material vor der Sendung vorbereiten, die konkreten Wetteraussichten sind allerdings live. In anderen Fällen ist der Wetterbericht schon vorher aufgezeichnet. Somit können manche Wettersendungen im semi-live Modus gesendet werden, sofern die Respeaker\*innen ausreichend zeitliche Ressourcen für die Vorbereitung der (oft recht kurzfristig gelieferten) Wetter-Endfassung haben.

In der gegenständigen Analyse wurden die Untertitel von 189 Sendungen vorproduziert oder konnten im semi-live Modus gesendet werden. Dies betraf hauptsächlich Nachrichtensendungen und Wetterberichte (182 Sendungen) mit 96 Prozent. Den restlichen Anteil machten Magazinsendungen (sieben Sendungen) mit vier Prozentpunkten aus. Musiksendungen, Dokumentationen und Quizsendungen wurden nicht in diesem Modus bearbeitet (siehe Abb. 12).

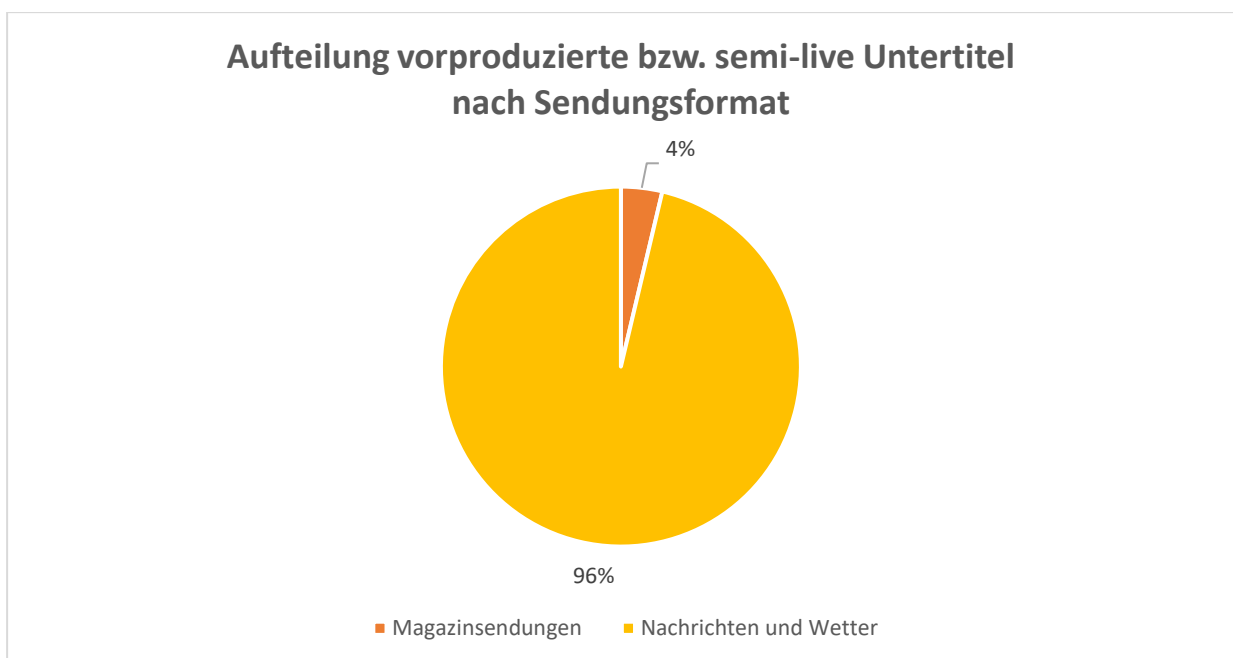


Abbildung 12: Aufteilung vorproduzierte bzw. semi-live Untertitel nach Sendungsformat in Prozent

Zu den Nachrichtensendungen sei angemerkt, dass üblicherweise Beitragstexte inklusive Videos, An- bzw. Abmoderation und gar Interviews bzw. Schaltungen schon vor Beginn der Sendung vorliegen. Daher lassen sich diese bestenfalls – falls diese vorbereitet werden können – im semi-live Modus untertiteln. Es kann jedoch vorkommen, dass Texte noch kurz vor oder gar während der Sendung bearbeitet oder verändert werden bzw. neu hinzukommen. Die Live-Untertitler\*innen müssen darauf dementsprechend reagieren und eventuell Texte live einsprechen. Auch Interviews bzw. Schaltungen müssen teilweise live untertitelt werden, weil sie entweder nicht voraufgezeichnet sind oder das Video nicht rechtzeitig vorbereitet werden konnte. Da diese Fälle aber eher die Ausnahme darstellen, sind die Untertitel von Nachrichtensendungen als semi-live zu kategorisieren.

Reine live-produzierte Untertitel lassen sich bei einer Dokumentationssendung (13 Prozent), fünf Magazinsendungen (63 Prozent) und zwei Musiksendungen (25 Prozent) finden. Nachrichten inklusive Wetter und Quizsendungen sind in dieser Kategorie nicht vertreten (siehe Abb. 13).

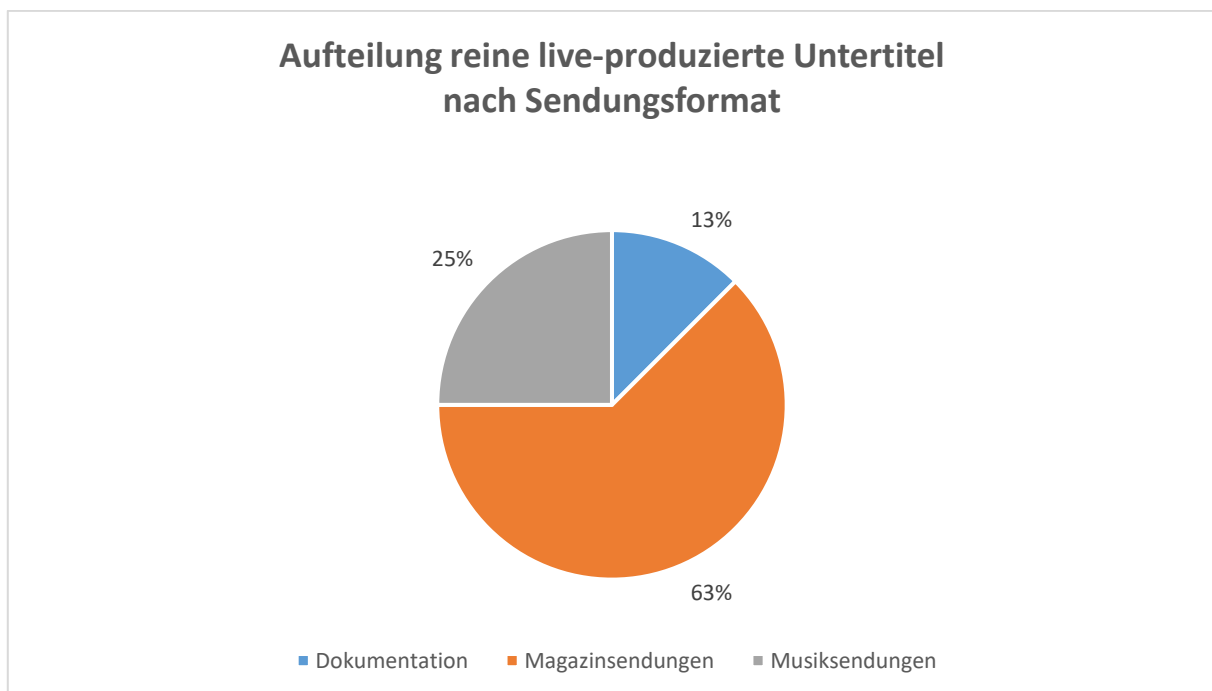


Abbildung 13: Aufteilung reine live-produzierte Untertitel nach Sendungsformat in Prozent.

Bei den reinen live-produzierten Untertiteln handelt es sich um voraufgezeichnete Sendungen, die nur sehr kurzfristig zur Verfügung gestellt wurden, oder um Sendungen, die komplett live ausgestrahlt wurden und bei denen MAZen oder An- bzw. Abmoderationen aufgrund von nicht verfügbarem Vorbereitungsmaterial nicht vorbereitet werden konnten.

Generell lässt sich aus der Aufstellung ablesen, dass Informationssendungen zum Großteil (72 Prozent) im semi-live Modus bearbeitet wurden. In knapp über einem Viertel der Fälle wurden gemischte Untertitel (27 Prozent) gesendet. Reine live-produzierte Untertitel kamen in dieser Programmsparte so gut wie nie vor. Anders sieht dies bei nonfiktionalen Unterhaltungssendungen und Sportsendungen aus. Hier liegt der Prozentsatz von reinen live-produzierten Untertiteln bei 67 Prozent bei nonfiktionalen Unterhaltungssendungen bzw. 40 Prozent bei Sportsendungen. Auch gemischte Untertitel lassen sich in diesen beiden Programmsparten häufiger finden. Bei nonfiktionalen Unterhaltungssendungen sind es 33 Prozent, bei Sportsendungen 60 Prozent. Im semi-live Modus wurden diese Programmsparten nicht untertitelt. Dies lässt sich damit begründen, dass in solchen Sendungen vermehrt Gäste eingeladen werden und die Gespräche mittels Respeaking live untertitelt werden müssen.

Je nach Programmsparte bzw. Sendungsformat sieht der Workflow für die Bearbeitung einer Sendung unterschiedlich aus. Im nächsten Unterkapitel soll versucht werden, den Arbeitsablauf bei der Live-Untertitelung zu veranschaulichen.

## **5.2. Workflow**

Allgemein kann der Arbeitsablauf bei der Live-Untertitelung in Vorbereitung, Live-Teil und Nachbearbeitung einer Sendung unterteilt werden. Im gegenständlichen Kapitel wird der Workflow anhand einer Selbstreflexion beschrieben. Die Erkenntnisse werden mit den Ergebnissen von Interviews mit Respeaker\*innen, welche die gleiche Sendung bearbeiten, zusammengetragen, abgeglichen und ergänzt. Zur besseren Veranschaulichung der verwendeten Programme sind Screenshots eingefügt. Die Redaktionstexte wurden aus urheberrechtlichen Gründen allerdings unlesbar gemacht.

Für die Beschreibung des Workflows einer live-untertitelten Sendung, die ein\*e Respeaker\*in alleine untertitelt, wird eine 45-minütige Magazinsendung mit einem breit gefächerten Themenspektrum und gemischten Untertiteln herangezogen. Gegen Ende der Sendung gibt es ein Live-Interview mit einem oder mehreren Studiogästen. Die Sendung beinhaltet somit Aspekte, die sowohl bei reinen live-produzierten als auch bei vorproduzierten bzw. semi-live Untertiteln vorkommen. Diese Herangehensweise ermöglicht es, einen umfangreicheren Überblick über den Workflow bei der Live-Untertitelung zu verschaffen. Aufgrund der verschiedenen Sendungsformate, der Sendungslängen und der Schnelligkeit im (Live-)Fernsehen sei jedoch angemerkt, dass es immer wieder zu spontanen Änderungen im Ablauf und bei Texten kommen kann und somit selbst die gleiche Sendung, die etwa montags bis freitags ausgestrahlt

wird, verschieden bearbeitet werden muss. Die einzelnen Arbeitsschritte variieren demnach unterschiedlich stark und hängen besonders von vorhandenem Vorbereitungsmaterial ab. Das spiegelt sich auch bei der untersuchten Live-Sendung wider. Denn die Gespräche gegen Ende der Sendung sind teilweise aufgezeichnet, wodurch es möglich ist, die ganze Sendung im Vorfeld vorzubereiten und im semi-live Modus zu Untertiteln.

### **5.2.1. Vorbereitung**

Mit der Vorbereitung der Sendung beginnen die Respeaker\*innen rund 2,5 Stunden vor Sendungsbeginn. Erfahrungsgemäß bleibt so genug Zeit, die Inhalte vorzubereiten, das Spracherkennungsprofil zu trainieren und das Untertitelungsprogramm sendebereit zu machen. Die Vorbereitungszeit schwankt je nach Dauer und Aufbau einer Sendung sowie vorhandenem Vorbereitungsmaterial, wie IP1 erklärt:

Das ist sehr abhängig von der Sendung, grundsätzlich aber ist es davon abhängig, ob Sprechertexte zur Verfügung stehen, wie gut diese sind, ob sie auch schon O-Töne inkludieren oder nur den Sprechertext. Davon mache ich das abhängig, und die Erfahrung zeigt dann, wie lange man ungefähr brauchen kann. Ebenso ist relevant, ob Videomaterial zur Verfügung steht, mit dem man das vorab abgleichen kann. (Interview IP1:26-30)

Eine allzu frühe Vorbereitung ist meist jedoch nicht zielführend, weil Texte und MAZen weit im Voraus der Sendung noch nicht final sind und diese bei Änderungen wieder angepasst werden müssten. IP2 merkt an, dass die Vorbereitungsdauer ebenfalls von der Qualität der verfügbaren Texte abhängt, sowie wie viele Live-Teile in einer Sendung vorkommen. Wenn Sendungen komplett live sind, gibt es oft nur wenig konkretes Material für die Vorbereitung von semi-live Untertiteln. Daher ist die Vorbereitungszeit bei Sendungen mit vielen Respeaking-Teilen auch kürzer:

Es geht auch schneller, wenn die Texte, die ich von der Redaktion bekomme, besser sind. Andererseits gibt es Sendungen, die komplett live sind, und da habe ich oft wenig Information, die ich ausschreiben kann. Und dann ist die Vorbereitung eigentlich kürzer, weil ich nur Wörter ausschreibe und einspreche und keine Texte vorbereiten muss. (Interview IP2:155-158)

Zu Beginn der Arbeit ist sicherzustellen, dass die zu Untertitelnde Sendung im Redaktionssystem Open Media (siehe Abb. 14) mit dem Untertitelungsprogramm FAB verknüpft ist. So werden die Inhalte des Redaktionssystems wie Ablauf, Moderationstexte und Videos direkt in das Untertitelungsprogramm geladen und automatisch aktualisiert, falls es zu Änderungen kommt (siehe Abb. 15). Hierbei ist darauf zu achten, an welchem Standort die Sendung produziert wird und in welchem Redaktionssystem die jeweiligen Sendungsinhalte zu finden sind. In Abbildung 14 findet man dies etwa auf der linken Seite des Programms. In der Mitte ist der Sendungsablauf

aufgeführt. Klickt man auf einen der Einträge, öffnen sich rechts die entsprechenden Texte für die An- bzw. Abmoderation und die Beiträge.

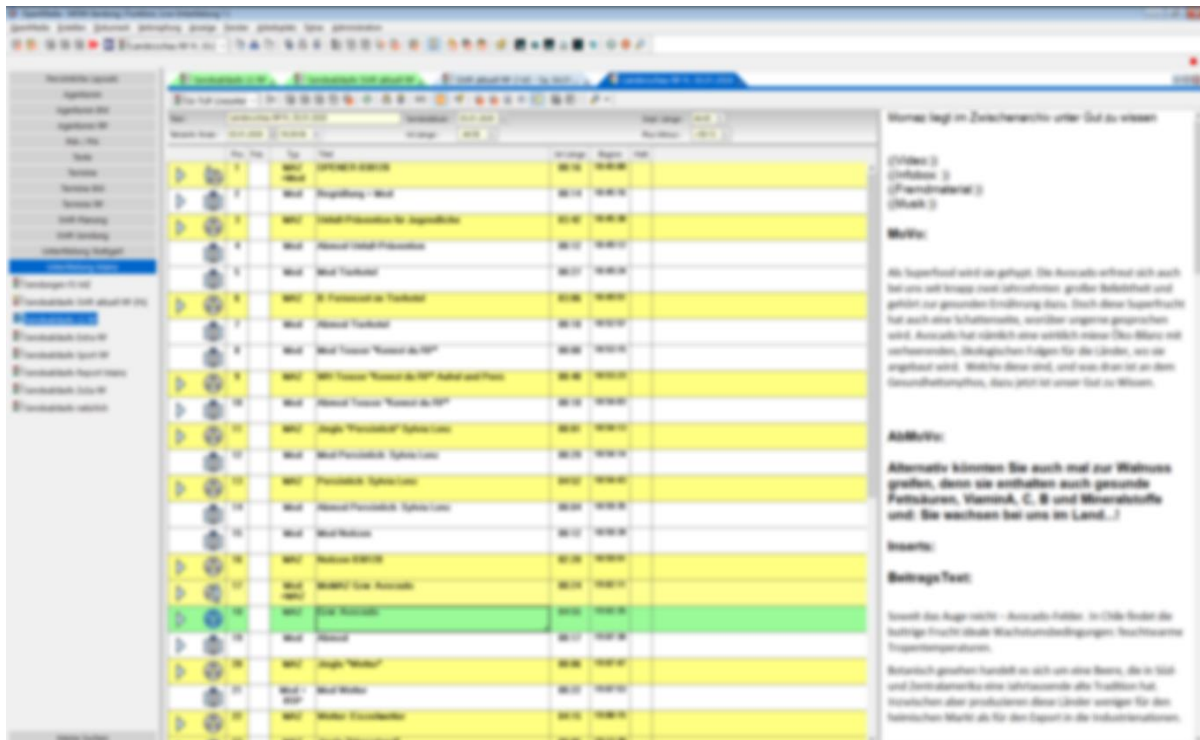


Abbildung 14: Screenshot des Programms Open Media und einer bestimmten Sendung

Sobald die Inhalte verknüpft sind, kann mit der Vorbereitung der vorhandenen Texte im Untertitelungsprogramm begonnen werden. Die Texte werden entsprechend den Untertitel-Richtlinien bearbeitet und gegebenenfalls gekürzt (siehe Abb. 15). Dabei lassen die Respeaker\*innen auch ihre Erfahrung miteinfließen:

Wenn man diverse Moderator\*innen kennt, kann man sich darauf einstellen, ob die Textqualität passt und man schon Sachen kürzen sollte, wenn man weiß, dass die Moderator\*innen schnell sprechen. (Interview IP2:236ff.)

Das Kürzen der Texte hilft dabei, dass die Zuseher\*innen genug Zeit zum Lesen der Untertitel haben. Auch IP1 verzichtet auf Füllwörter, die die Untertitel unnötig länger machen.

[I]n diesem Fall waren die Anmoderationen für die Beiträge vollständig. Ein paar Füllwörter wurden in der Live-Situation hinzugefügt. Die habe ich aber nicht beachtet, weil sie für den Inhalt nicht relevant waren. (Interview IP1:66ff)

Mit zunehmender Erfahrung hat sich herausgestellt, vorzugsweise mit dem Bearbeiten der Beitragstexte von MAZen zu beginnen, die bestenfalls schon abgenommen (das heißt sendefertig) sind, und An- bzw. Abmoderationen dahingegen erst kurz vor Beginn der Sendung vorzubereiten. Im Vorfeld gibt es oft nur Vorschläge für An- bzw. Abmoderationen, welche so auch im



Redaktionssystem vermerkt sind. Diese Texte sind noch nicht final und müssten bei Änderungen schlimmstenfalls komplett neu bearbeitet werden, was wiederum Vorbereitungszeit kostet, die anderweitig besser genutzt werden könnte. Gewisse Texte sind somit weit im Voraus vor Sendungsbeginn eher zu gebrauchen als andere. Diesen Aspekt sollten die Respeaker\*innen in der Vorbereitungsphase abwägen und berücksichtigen. Ob MAZen abgenommen sind, lässt sich im Redaktionssystem überprüfen.

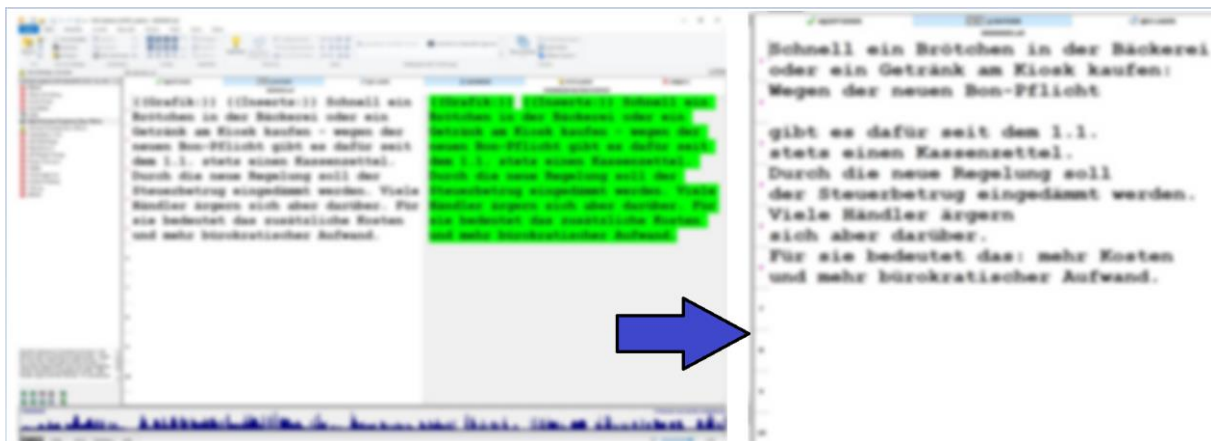


Abbildung 15: Screenshot des Untertitelungsprogramm FAB mit den geladenen Open Media-Inhalten

Sind die MAZen vor Sendungsbeginn vorhanden, ist es außerdem sinnvoll, sofern genügend Zeit zur Verfügung steht, die Videos mit dem Beitragstext abzugleichen, um etwaige Unterschiede von Text und Bild oder fehlende Textpassagen ausfindig zu machen. Manchmal kommt es vor, dass O-Töne<sup>10</sup> oder Voxpops<sup>11</sup> im Beitragstext nicht vorhanden sind und diese dann während der Sendung live gerespeakt werden müssen, falls man diese (zumeist nur sehr kurzen) Teile nicht vorbereiten konnte. IP2 beschreibt den Ablauf wie folgt:

Dann habe ich mir eben im System angesehen, welche Beiträge dieser Sendung schon fertig sind und diese zuerst vorbereitet. Meistens schaue ich einfach, wo das Video verfügbar ist, und kontrolliere dann die Texte, die mir zur Verfügung stehen seitens der Redaktion gleich mit dem Video, während ich das Ganze vorbereite, damit etwaige O-Töne – die stellen meistens das größte Problem dar – dann auch schon stimmen in der Live-Situation. (Interview IP2:120-124)

Durch das vorherige Abgleichen kann der Beitragstext dem sendefertigen Video angepasst und O-Töne eingefärbt werden, damit die Untertitel während der Live-Sendung vollständig sind.

<sup>10</sup> O-Töne steht für "Originaltöne". Dabei handelt es sich um Wortmeldungen bzw. Statements von Interviewpartner\*innen. Zumeist sind dies interviewte Expert\*innen zu einem Thema oder Reporter\*innen vor Ort.

<sup>11</sup> Voxpops sind Wortmeldungen von befragten Personen auf der Straße – meistens ohne Inserts, weil oft nicht mehr als ein bis zwei Sätze gesprochen werden.

Außerdem könnte es vorkommen, dass zwar Videos vorhanden sind, aber bis kurz vor Sendungsbeginn kein Beitragstext vorliegt. Den Text können die Respeaker\*innen mithilfe der Spracherkennungssoftware dann relativ flott einsprechen:

Es gab aber auch einen Teil, wo nur ein Video verfügbar war, wo ich selbst den Text mittels Spracherkennung hinzugefügt habe. (Interview IP1:6ff)

Es sei aber angemerkt, dass die Bearbeitung mit vorhandenen Texten schneller vonstattengeht, als wenn der Text erst selbst erstellt werden muss. Nach diesen Schritten verfährt man weiter mit den vorhandenen Beitragstexten, Videos und An- bzw. Abmoderationen, bis alle abgearbeitet sind.

Da für das Ende der hier beleuchteten Sendung noch ein Live-Gespräch mit Studiogästen geplant war, mussten sich die Respeaker\*innen dementsprechend auf das Interview vorbereiten. Gästenamen und Fragen der Moderator\*innen können im Open Media vorher angesehen werden. Manchmal sind sogar stichpunktartige Antworten bei den Fragen dabei. Dadurch kann das Ziel des Interviews abgeschätzt und ein roter Faden erkennbar werden. Mithilfe des Materials können die Respeaker\*innen Fachtermini und Eigennamen im Redaktionssystem recherchieren und diese dann der Spracherkennungssoftware beibringen. Zudem lohnt es sich, Namen von Gästen und Orten oder sonstige Termini auf ein Blatt Papier zu schreiben, damit man im Eifer des Gefechts die richtige Schreibweise vor Augen hat, falls die Spracherkennung die Ausdrücke nicht erkennt. Besondere Vorsicht ist bei Wortspielen und Eigennamen geboten. Manche Moderator\*innen schreiben sich Namen in der Lautschrift auf, damit sie ihn auch richtig aussprechen. Die Schreibweise sollte man nochmal bei den vorhandenen Informationen im Open Media oder gegebenenfalls im Internet überprüfen.

Um die Liste der Wörter dem Spracherkennungssystem beizubringen oder eventuelle Videos ohne Beitragstext vorzubereiten, muss zuerst die Software Dragon Naturally Speaking mit dem dazugehörigen Profil der Respeaker\*innen geöffnet werden. IP2 merkt an, dies schon zu Beginn des Arbeitsablaufs zu erledigen, weil es einige Zeit dauern kann, bis das Profil – das mit dem Training immer größer wird – einsatzfähig ist.

Dann habe ich mein Sprachprofil für die Spracherkennung kopiert. Das dauert nämlich immer einige Zeit, weil es schon recht groß ist. Deswegen versuche ich das immer ganz am Anfang zu machen, damit ich – während alle anderen Systeme laden – das schon kopiere und es bereitsteht, spätestens wenn die Sendung beginnt. (Interview IP2:116-120)

Für die untersuchte Sendung recherchierte der Verfasser der Masterarbeit 83 Wörter und schrieb diese in einer Liste nieder. Davon wurden jedoch nur 33 Vokabel in das Spracherken-

nungssystem eingespeist (siehe Abb. 16). Die restlichen Wörter waren der Software schon bekannt. IP1 hat ihrem Sprachprofil elf Vokabel beigebracht. Den Fokus legte sie dabei auf Eigennamen (mit Titel bzw. Anrede) sowie auf Termini wie medizinische oder geographische Bezeichnungen. Das sind vor allem jene Ausdrücke, bei denen nicht davon auszugehen ist, dass die Spracherkennung diese Wörter ohne vorheriges Trainieren erkennen bzw. verstehen wird. Welche Wörter sie einspeichert, macht sie von ihrer Erfahrung abhängig:

Ich kenne mein Spracherkennungsprogramm schon recht gut und daher weiß ich, dass manche Dinge sicherlich gut erkannt werden, und habe mich auf Eigennamen fokussiert, sowohl mit Titel bzw. mit Anrede, also „Frau etc.“ oder „Mann etc.“ Einmal mit Vor- und Nachnamen. Dann kam auch ein Hund vor. Dessen Namen habe ich auch ausgesprochen, und ein paar geographische Begriffe waren dabei, wo ich nicht sicher war, ob es gut erkannt werden würde, und zwei medizinische Begriffe ebenfalls. In dem Fall waren es also nur Nomen, die ich ausgesprochen habe. (Interview IP1:17-23)

IP2 hat rund 20 bis 30 Wörter vorbereitet. Sie sprach aber nicht alle ein, weil die anderen schon im Sprachprofil vorhanden waren. Zudem gab sie an, dass im Live-Gespräch auf die aktuelle politische und gesundheitliche Lage eingegangen und somit vieles sehr allgemein gehalten wurde. Da damit in Verbindung stehende Ausdrücke im Fernsehen des Öfteren vorkommen, hatte sie diese Vokabel ihrem Sprachprofil schon für andere Sendungen beigebracht. Auch sie schätzt im Vorhinein ab, mit welchen Wörtern das Spracherkennungssystem eventuell Probleme haben könnte, und fügt ihm diese hinzu:

Das waren unterschiedliche Wörter, wie Nomen und Verben, bei denen ich davon ausgehe, dass die Spracherkennung sie nicht kennt. So etwas wie „Atomkrieg“, zum Beispiel, habe ich in meiner Liste gesehen, aber auch Eigennamen von einer Band oder der Name des Gastes am nächsten Tag, natürlich auch der Name des Gastes selber und seiner Rolle, die er spielt in seinem Kabarettprogramm. Hauptsächlich handelt es sich um Eigennamen, aber auch alle Verben und Nomen, die nicht dem Standardrepertoire der Spracherkennung angehören, wobei das im Prinzip in meinem eigenen Ermessen liegt, wie ich die einschätze, was ich glaube, was die Spracherkennung nicht kann. (Interview IP2:138-145)

Die Respeaker\*innen sollten ihr Sprachprofil demnach gut kennen, um zu wissen, welche Termini sie in der Vorbereitungsphase ihrem Profil einzuspeisen haben, damit sie in der Live-Situation davon ausgehen können, dass das Spracherkennungssystem diese Ausdrücke auch fehlerfrei erkennt.

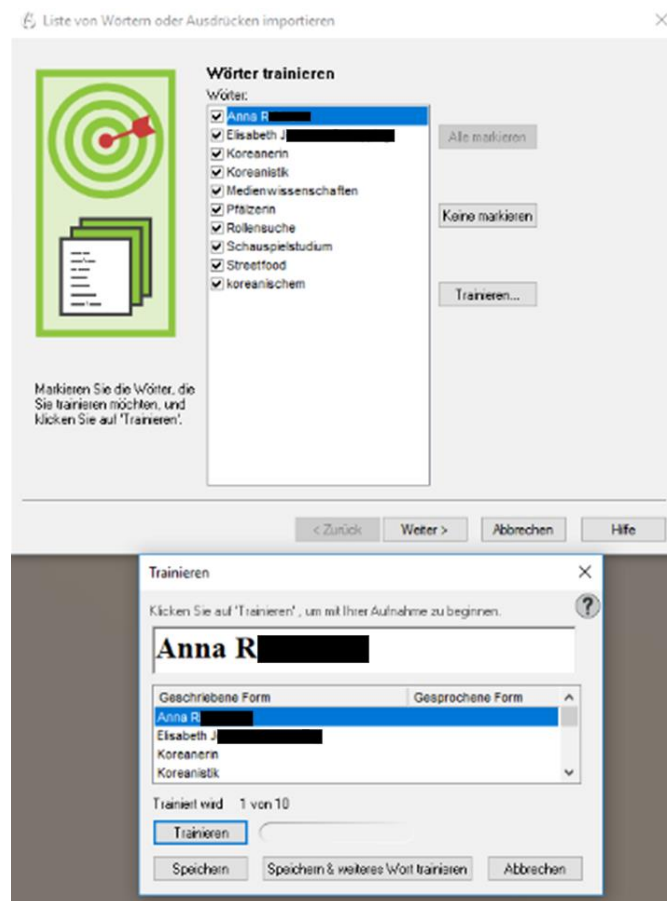


Abbildung 16: Screenshot des Spracherkennungsprogramms beim Hinzufügen und Eintrainieren unbekannter Wörter

Vor Sendungsbeginn sollte noch einmal geengeprüft werden, ob die Mikrofontechnik reibungslos funktioniert, die Spracherkennungssoftware die Respeaker\*innen gut erkennt und die trainierten Wörter auch korrekt wiedergibt.

Zudem gibt es für kurz vor Sendungsbeginn eine Checkliste, an die sich die Respeaker\*innen halten können, um sicher zu gehen, dass die Untertitel auch beim richtigen Sender ankommen. Das jeweilige Rückbild und das Live-Bild werden mithilfe eines Video Control Panels eingestellt. Welcher Sendeweg bzw. welche Sendewege in der Untertitelungssoftware auszuwählen sind, ist unter anderem im Dienstplan vermerkt. Ob die Untertitel im Fernsehen angezeigt werden, können die Respeaker\*innen mittels Rückbild eines kleinen Bildschirms überprüfen, der in der Kabine angebracht ist. Allerdings funktioniert dies erst ab dem Moment, wenn man live auf Sendung ist. Außerdem muss an der Sprechstelle überprüft werden, ob auch der richtige Ton des jeweiligen Senders empfangen wird. Wenige Sekunden vor Beginn der Sendung wird dann im Untertitelungsprogramm der Live-Modus aktiviert und die Respeaker\*innen können auf Sendung gehen. Erst danach können Untertitel geschickt und im Fernsehen angezeigt werden. Es ist auffällig, dass beide interviewten Respeaker\*innen keinen dieser Punkte erwähnt haben. Es kann davon ausgegangen werden, dass sie diese Schritte schon stark

verinnerlicht haben, weil ohne diese Einstellungen auch keine Untertitel bei den Sendeanstalten ankommen würden.

### **5.2.2. Live-Sendung**

Sobald die Sendung startet, senden die Respeaker\*innen den ersten Untertitel. Je nach Sendeanstalt gibt es verschiedene Richtlinien, wie der erste Untertitel auszusehen hat und ob anzugeben ist, dass es sich bei dem Programm um Live-Untertitel handelt. Direkt nach Beginn der Sendung sollte am Rückbild kontrolliert werden, ob die gesendeten Untertitel aufscheinen.

Da bei den gegenständlich analysierten Sendungen (fast) alle MAZEn vor Sendungsbeginn auf inhaltliche Richtigkeit und Vollständigkeit hin überprüft werden konnten und die vorhandenen An- bzw. Abmoderationen von dem\*der Moderator\*in eingehalten wurden, konnten die vorbereiteten Untertitel im semi-live Modus abgeschickt werden (siehe Abb. 17). IP1 erwähnt jedoch, dass sie während eines Beitrags, den sie vorab nicht kontrollieren konnte, fehlende O-Töne live einsprechen musste.

Bei dieser spezifischen Sendung heute gab es einen O-Ton, der nicht markiert war in einem Beitrag. Das waren ungefähr drei Sätze, wo ich dann umschalten musste auf ein „Sprachinterface“, wo ich dann direkt einsprechen kann und meinen Untertitel hinausschicken kann. (Interview IP1:59-62)

Dies hat allerdings zur Folge, dass darunter die Synchronität von Bild und Ton leidet. Zudem müssen die Respeaker\*innen manche O-Töne mit dem Drücken einer Taste (in dem Fall sind es die F-Tasten auf der Tastatur) live einfärben, falls dies nicht in der Vorbereitung gemacht werden konnte. Des Weiteren überprüfen sie die Untertitel während des Sendens auf orthografische und grammatikalische Richtigkeit. Etwaige Fehler versuchen sie noch vor dem Senden auszubessern:

Wenn ein Fehler im Untertitel, der vorbereitet war, aufkommt, versucht man diesen noch schnell zu korrigieren, bevor man ihn mit Enter rausschickt. (Interview IP1:56f.)

IP1 fügt hinzu, dass nicht immer alle sprachlichen Mängel korrigiert werden können und Fehler bei gesendeten Untertiteln nicht komplett vermeidbar sind. Sie notiert sich daher eine kurze Wortfolge, um den Fehler mittels Suchfunktion bei der Nachbearbeitung schneller finden zu können:

[H]ier schreibe ich mir immer eine kurze Wortfolge auf, damit ich dann mittels Suchfunktion schneller an die Stellen komme, die nachbearbeitet werden müssen vom Textinhalt. (Interview IP1:79ff.)

IP2 geht bei auftretenden Fehlern ähnlich vor:

Im Idealfall kontrolliere ich auch live noch, ob Fehler drinnen sind, was sehr häufig auch noch der Fall ist, weil ich irgendetwas übersehen habe. Und dann bessere ich das schnell aus, bevor ich es wegschicke. (Interview IP2:175ff.)

Bei der Sendung heute habe ich schon währenddessen einige Fehler ausbessern können, sodass ich mir die nicht aufschreiben musste. (Interview IP2:206f.)

Während des Live-Teils bedienen die Respeaker\*innen außerdem verschiedene Funktionen im Untertitelungsprogramm. In Abbildung 17 sind links die jeweiligen Beiträge aufgelistet, die dem Sendeablauf entsprechen. Rechts werden die vorbereiteten Untertitel angezeigt. In der Mitte ist das Sendebild eingeblendet. Darin sieht man den aktuellen Untertitel, der gerade gesendet wird. Dabei gilt es, die sogenannte Standzeit einzuhalten, damit die Zuseher\*innen genügend Zeit haben, den Untertitel zu lesen. Sobald der schwarze Balken vom roten Bereich in den grünen übergeht, ist die Einhaltung der Standzeit gegeben. Unter dem Balken ist der nächste Untertitel eingeblendet, der gesendet wird, sobald die Respeaker\*innen ihn mit Enter abschicken.

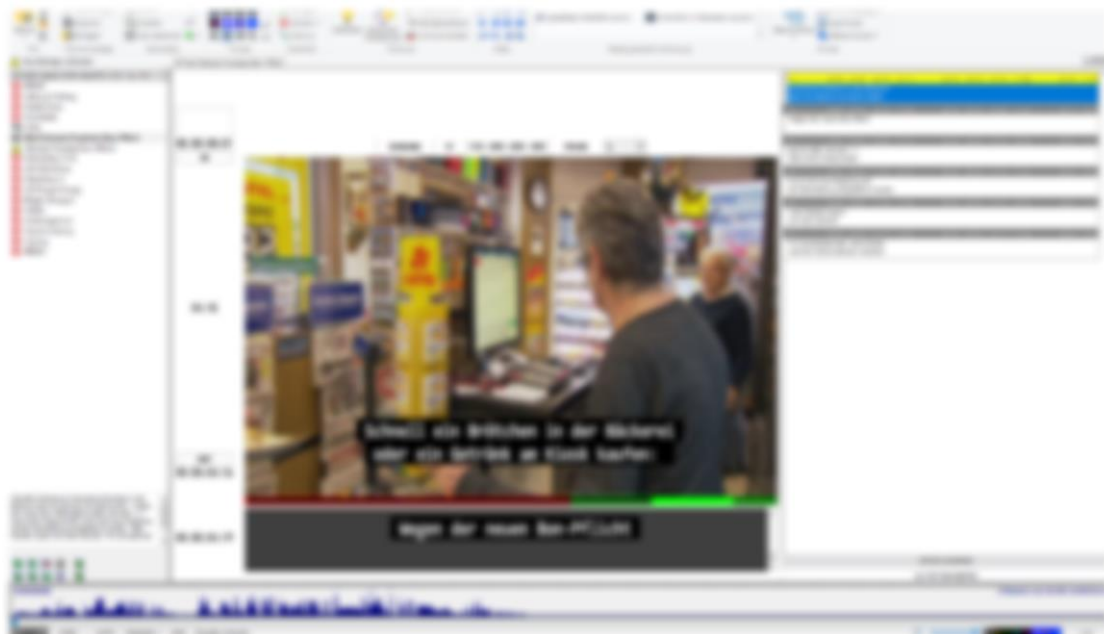


Abbildung 17: Screenshot beim Live-Untertiteln im semi-live Modus

Die Untertitel werden standardmäßig am unteren Rand angezeigt. Da jedoch Inserts wie etwa von O-Tönen auch unten eingeblendet werden, müssen die Respeaker\*innen die Untertitelposition während der Live-Sendung mit einer Tastenkombination entsprechend anpassen, um die Inserts nicht zu überdecken:

Und ich muss bei Inserts, bei Namen oder Reporter-Inserts, die Untertitel hochstellen, damit man das dann im Fernsehen trotzdem noch alles lesen kann. (Interview IP2:173ff.)

Da die Inserts in den Vorbereitungsvideos noch nicht vorhanden sind, müssen die Respeaker\*innen die Untertitelposition live ändern. Wenn durch das Hochstellen der Untertitel das

Mundbild der Sprecher\*innen verdeckt werden würde, könnten die Untertitel auch am oberen Rand platziert werden. Dadurch ist es gehörlosen Menschen und Menschen mit Hörbeeinträchtigung möglich, weiterhin das Lippenlesen zum besseren Verständnis zu nutzen. Erfahrungsgemäß ist es zudem sinnvoll, bei aufkommenden O-Tönen die Untertitelposition jedenfalls anzupassen – auch wenn das Insert noch nicht angezeigt wird oder es teilweise zu gar keinem kommt. Grund dafür ist eine kleine Verzögerung beim Senden und Anzeigen der Untertitel im Fernsehen. Oft werden die Untertitel zwar hochgestellt, aber durch die Verzögerung kann es trotzdem zu einer Überdeckung des Inserts kommen. Werden die Untertitel schon im Vorhinein hochgestellt, bevor das Insert überhaupt sichtbar ist, kann man dieses Problem umgehen.

Beim Interview mit den Studiogästen kam dann die Respeaking-Methode zum Einsatz. Dieser Teil dauerte beim Verfasser der Masterarbeit etwa neun Minuten, bei IP2 rund zehn bis 20 Minuten. IP1 konnte dank der Voraufzeichnung das Gespräch schon vor Sendungsbeginn transkribieren und im semi-live Modus abschicken:

Das Video kam vorab, das heißt, ich habe diesen Teil nicht live-live gemacht, sondern im Endeffekt vorbereiten können. Das habe ich auch mit Spracherkennung gemacht. (Interview IP1:10ff.)

Im Live-Gespräch schalten die Respeaker\*innen mit einer Tastenkombination ins sogenannte Sprachinterface. Dabei öffnet sich im Untertitelungsprogramm ein Diktierfenster unterhalb des Sendebildes. In dieses Fenster diktieren die Respeaker\*innen den Text (siehe Abb. 18). Zudem muss noch die Spracherkennungssoftware mit einem Tastenklick, den man individuell festlegt, aktiviert werden. Danach kann mit dem Respeaking begonnen werden. Der diktierte Text wird zuerst im Textfeld angezeigt. Es ist möglich, diesen noch vor dem Abschicken auszubessern und einzufärben. Auch die Untertitelposition können die Respeaker\*innen anpassen. Erst nach dem Abschicken ist der Untertitel im Fernsehen sichtbar. Hierbei sollte darauf geachtet werden, die Untertitelkonventionen möglichst einzuhalten, damit die Untertitel gut lesbar sind. Obwohl im gegenständlichen Fall die Interviewfragen ident mit den Fragen im Redaktionssystem waren, lohnt es sich jedoch kaum, für die kurzen Fragen vom Diktiermodus in den semi-live Modus zu schalten, weil dies zu viel Zeit gekostet hätte und eine größere Lücke bei den Untertiteln entstanden wäre. Praktischer ist es, die Fragen auch zu respeaken, damit der Untertitelfluss konstant bleibt.



Abbildung 18: Screenshot beim Live-Untertiteln im Respeaking-Modus

Nach dem Interview verabschiedete sich der\*die Moderator\*in und übergab zur nächsten Sendung. Auch diesen Abschnitt konnten die Respeaker\*innen wieder im semi-live Modus untertiteln. Allerdings ist hier anzumerken, dass die Live-Interviews nicht immer sekundengenau eingehalten werden und dadurch das Ende der Sendung schnell eingeläutet werden kann. Dabei improvisieren die Moderator\*innen oftmals oder die Sendeanstalt macht einen regelrechten Cut und schaltet zur nächsten Sendung. Die Respeaker\*innen müssen daher sehr flexibel auf etwaige Änderungen oder Kürzungen reagieren. Auch aus diesem Grund sind voraufgezeichnete Interviews von Vorteil, weil die Sendung so ohne Zeitdruck beendet werden kann. Nach Ende der Sendung wird von Sendung gegangen, damit keine weiteren Untertitel die Leitung „blockieren“ können.

### 5.2.3. Nachbearbeitung

Da die vorliegende Sendung wiederholt wurde, musste sie dementsprechend für die Fernseh-wiederholung, die Online-Mediathek und das Archiv überarbeitet werden. Ob Sendungen noch-mal ausgestrahlt werden, lässt sich aus dem Dienstplan ablesen. Die fertige Untertiteldatei muss dann in ein Video-System hochgeladen und mit dem Sendungsvideo verknüpft werden.

Die Nachbearbeitung beinhaltet vor allem sprachliche Korrekturen, inhaltliche Ergänzungen, Interpunktion sowie das Anpassen der Untertitelpositionen und der Timecodes für genauere Synchronität. Hauptsächlich wurden die Respeaking-Teile für die untersuchte Sendung



überarbeitet, weil die semi-live Teile aufgrund der Vorbereitung weniger fehleranfällig waren. IP2 beschreibt ihre Nachbearbeitung wie folgt:

[E]s [sind] oft Beistrichfehler oder große oder kleine Anfangsbuchstaben, die ich gerne noch ausbesser. Dann muss ich darauf achten, dass jeweils beim Wechsel vom voraufgezeichneten Video zum Moderator der erste Untertitel immer genau am Anfang der Tonspur ist. Und beim Live-Teil habe ich natürlich mehr Fehler ausbessern müssen, einige Sachen hinzugefügt oder korrigiert, wenn ich etwas nicht ganz richtig verstanden habe. (Interview IP2:208-212)

Generell kann festgehalten werden: Je weniger Vorbereitung und/oder je umfassender die Respeaking-Teile, desto länger und umständlicher ist die Nachbearbeitung einer Sendung. Dafür muss zuerst das Sendungsvideo von einem Video-System heruntergeladen werden. Außerdem werden die gesendeten Untertitel in einer Datei abgespeichert. Danach kann man das Video sowie die Untertiteldatei in das Untertitelungsprogramm laden und die Korrekturen und Anpassungen durchführen. Dabei gilt es darauf zu achten, die Timecodes dem Video anzupassen und den ersten Timecode beim Programmstart anzugeben. Dies geschieht, damit die Untertitel-Timecodes mit den Zeiten des Videos übereinstimmen und die Untertitel synchron zu Bild und Ton laufen:

Dann setze ich den ersten Untertitel mit dem richtigen Timecode und kontrolliere nochmal alle Timecodes, alle Untertitel auf die Standzeit, ob die auch korrekt ist und ob der Untertitel an der richtigen Stelle erscheint. (Interview IP1:82ff.)

Sobald die Sendung überarbeitet ist, muss die Untertiteldatei noch passend benannt und in einem eigenen Ordner abgespeichert werden. Die Vorgabe zur Benennung stammt meist von den Sendeanstalten selbst. Anschließend wird die Datei im Video-System mit dem sendefertigen Video verknüpft. Dabei kann noch überprüft werden, ob die Untertitel mit dem Beitragstext des Videos synchron übereinstimmen. Abschließend wird die Datei ans sogenannte Playoutcenter geschickt, damit sie für die TV-Wiederholung bei der jeweiligen Sendezentrale vorliegt. Vorsichtshalber wird die Datei auch noch per E-Mail versendet.

Nachdem die fertige Datei abgespeichert und abgeschickt worden ist, lohnt es sich, falsch erkannte Wörter dem Spracherkennungssystem beizubringen und es darauf zu trainieren. Dieser Schritt wurde in der gegenständlichen Sendung aus Zeitgründen jedoch nicht durchgeführt. Zu guter Letzt wird das Sprachprofil abgespeichert und der Computer kann danach neu gestartet werden.

### 5.3. Problemfaktoren bei der Live-Untertitelung

Während der einzelnen Phasen sind die Respeaker\*innen mit verschiedenen Problemfaktoren konfrontiert. Diese werden nun erläutert und ebenfalls in Vorbereitung, Live-Sendung und Nachbearbeitung aufgegliedert.

#### 5.3.1. Vorbereitung

In der Vorbereitungsphase sind vor allem vorhandene Texte und Videos von großer Bedeutung. Sind diese fehlerhaft, unvollständig oder gar nicht vorhanden, erschwert dies die Arbeit der Respeaker\*innen. Besonders O-Töne stellen einen Problemfaktor dar. Diese fehlen in den Beitragstexten teilweise ganz und/oder können nicht mit der dazugehörigen MAZ gegenkontrolliert werden. Das kann in der Live-Situation zu Herausforderungen führen. Mit diesen Problemfaktoren müssen die Respeaker\*innen laut IP2 bei so gut wie jeder Sendung rechnen. Sie geht davon aus, dass das Redaktionsteam nicht viel Wert auf grammatikalisch korrekte Texte und saubere Interpunktion legt. Diese Annahme bestätigt auch Käber (2016), der Live-Untertitler bei der ARD Gemeinschaftseinrichtung ARD Text in Potsdam ist:

Bei vorbereiteten Untertiteln kann zwar auf die fertigen Beitragstexte im Redaktionssystem zurückgegriffen werden, allerdings kommt hier oft erschwerend hinzu, dass der Verfasser des Beitrags die Texte oft nur für die eigene Studioproduktion hinterlässt. Diese eher im Schnellverfahren verfassten Texte sind aber nicht zum Senden als geschriebener Untertiteltext bestimmt und müssen daher oft stark redigiert und von ungezählten Flüchtigkeitsfehlern befreit werden. (Käber 2016:304)

Für die Arbeit der Respeaker\*innen ist es jedoch überaus hilfreich, wenn Beitragstexte vollständig sowie fehlerfrei sind und/oder sie die Texte mit den Videos auf Vollständigkeit und Richtigkeit vor Sendungsbeginn abgleichen können. IP1 gibt des Weiteren an, dass Respeaker\*innen oft mehrere Sendungen gleichzeitig vorbereiten, sofern brauchbares Material verfügbar ist. Sie beschäftigen sich mit Beiträgen für andere Sendungen, was die Art der Vorbereitung für eine Sendung verändert. Gutes Zeitmanagement ist somit unabdingbar.

Ein weiterer Punkt, auf den die Respeaker\*innen zu achten haben, ist die Genauigkeit bei vorhandenen Texten. IP2 erwähnt, dass sie ein Wortspiel nicht erkannt und es daher ihrem Sprachprofil falsch beigebracht habe:

Heute in der Sendung war ein Kabarettist zu Gast. Dann habe ich mir auch den Namen seines aktuellen Programms bzw. seiner neuen Sendung eingespeichert, habe aber offensichtlich nicht genau genug gelesen. Denn der Titel ist ein Wortspiel und ich habe das eigentliche Wort genommen, und nicht so, wie die Sendung heißt, nämlich das Wortspiel. [...] Da hätte ich mich gleich besser vorbereiten sollen und genauer lesen sollen. (Interview IP2:126-131)

Den Fehler musste IP2 während der Sendung live korrigieren, was möglicherweise zu Verzögerungen beim Senden der Untertitel bzw. einem kognitiven Mehraufwand geführt hat. Problemfaktoren können somit auch Anspielungen und Wortspiele sein, die in der Vorbereitungsphase nicht erkannt werden.

### **5.3.2. Live-Teil**

Die Untersuchung zeigt auf, dass die Respeaker\*innen im Live-Teil sehr häufig mit Technikproblemen zurechtkommen müssen. Diese können das Ausfallen von Bild und Ton betreffen sowie Probleme mit den verwendeten Programmen für das Erstellen der Untertitel bzw. mit der Spracherkennungssoftware:

Technikprobleme, sowohl Ton als auch Bild können ausfallen, das Programm kann zu langsam sein, sowohl das Untertitelprogramm als auch das Spracherkennungsprogramm. Das sind die größten Probleme, die auftreten. (Interview IP1:70-73)

IP2 berichtet von konkreten technischen Schwierigkeiten bei der untertitelten Sendung:

Ja, das war auch heute der Fall, dass sich plötzlich meine Ansicht am PC verkleinert hat und ich alle offenen Programme am PC gesehen habe. Dadurch hat natürlich die Spracherkennung nicht in mein Untertitelprogramm gesprochen. Dann musste ich manuell wieder reinklicken und alles wiederholen. (Interview IP2:183-186)

Diese Problemfaktoren können sowohl beim Untertiteldienstleister auftreten als auch bei der Sendeanstalt. Verzögerungen bei den gesendeten Untertiteln sind wahrscheinliche Konsequenzen.

Eine weitere Herausforderung liegt in der Respeaking-Methode. Zum einen kann es wegen der Spracherkennungssoftware zu Problemen kommen, wenn diese die diktierten Wörter nicht genau genug oder falsch erkennt oder einzelne Ausdrücke überhaupt nicht transkribiert. Die Mängel haben die Respeaker\*innen im Nachhinein selbst zu korrigieren. Zum anderen betrifft es die zu untertitelnden Personen. Schnelle Sprechgeschwindigkeit, Dialekte, Sprachvarietäten und Witze stellen Problemfaktoren dar. Diese können je nach Sendung unterschiedlich sein, jedoch sind sie laut IP2 in fast jeder Sendung anzutreffen:

Das kommt natürlich auf die Sendung an. Im Regionalfernsehen kommt mehr Dialekt vor als im überregionalen Fernsehen. Da ist die Sprachfärbung nicht so das Problem. Dafür gibt es im Entertainmentbereich, was überregional ist, oft schnelle Sprache, Umgangssprache, Anglizismen. Aber im Endeffekt ist das eine oder andere Problem in jeder Sendung zu finden. Die Frage ist, ob man sich davon aus der Ruhe bringen lässt oder nicht. (Interview IP2:198-203)

Von den Respeaker\*innen erfordert es demnach passende Strategien, um mit diesen Herausforderungen umzugehen. Jedoch ist IP1 der Meinung, dass die Problemfaktoren weniger in der Sprachdienstleistung an sich, sondern eher in der Technik zu finden sind.

Während der Sendung kommt es zudem manchmal zu Änderungen im Programmablauf und Beiträge werden in ihrer Reihenfolge vertauscht. Dies kann laut IP1 zu Verwirrung führen. Die Respeaker\*innen müssen deswegen auch eine gewisse Flexibilität während der Live-Situation an den Tag legen und bei Änderungen beachten, welche Texte sie vorbereitet haben, die nur durcheinandergebracht wurden, oder welche Texte sie live zu respeaken haben.

### **5.3.3. Nachbearbeitung**

Die interviewten Respeaker\*innen berichten auch in der Nachbearbeitungsphase besonders häufig von Technikproblemen, mit denen sie sich auseinandersetzen müssen. Sendungsmitschnitte sind eventuell nicht bzw. relativ spät verfügbar oder unvollständig. Außerdem kann die nachzubearbeitende Untertiteldatei Fehler aufweisen. Des Weiteren sind Probleme beim Untertitelungsprogramm festzustellen, wenn dieses nicht reibungslos läuft. Aufgrund dieser Faktoren müssen die Respeaker\*innen mehr Zeit für die Nachbearbeitung aufwenden. Im Allgemeinen korreliert die Nachbearbeitung sehr stark mit der Vorbereitung und der Live-Sendung, wie IP2 anführt:

Je besser vorbereitet ist und je besser es live läuft, desto unproblematischer läuft die Nachbearbeitung. Je ungenauer man davor gearbeitet hat, desto umständlicher und länger wird die Nachbearbeitung. (Interview IP2:224ff.)

Gibt es kein oder nur wenig brauchbares Vorbereitungsmaterial und/oder treten Problemfaktoren während des Live-Teils auf, müssen die Respeaker\*innen auch mehr Zeit in die Nachbearbeitung investieren.

### **5.3.4. Fazit**

Zusammenfassend lässt sich feststellen, dass die Respeaker\*innen während der drei Phasen in der Live-Untertitelung mit mehreren Problemfaktoren konfrontiert sein können. Diese sind von Sendung zu Sendung verschieden und unterscheiden sich gar bei den gleichen Sendungen angesichts neuer Beiträge, wechselnder Themenbereiche, Gäste und Moderator\*innen. Die Erfahrung hilft, Problemfaktoren schon vor der Sendung zu erkennen und sich entsprechende Strategien zurechtzulegen.

Des Weiteren zeigt sich, dass gut aufbereitetes Vorbereitungsmaterial die Arbeit der Respeaker\*innen deutlich erleichtert. Je mehr brauchbares Material verfügbar ist, desto unkomplizierter verläuft die Live-Sendung und desto weniger muss nachbearbeitet werden. Daraus kann abgeleitet werden, dass auch eine bessere Qualität der Untertitel mit guter Vorbereitung einhergeht. Somit liegt es an den Sendeanstalten selbst, welches, wie viel und wie gut aufbereitetes Material sie zur Verfügung stellen, um so für bessere Untertitelqualität zu sorgen.

Zudem ist festzuhalten, dass es sich beim Respeaking um einen hochgradig technisch anspruchsvollen Beruf handelt. Die meisten Problemfaktoren lassen sich in der Technik finden. Die Respeaker\*innen müssen mehrere Programme (teilweise gleichzeitig) bedienen können, dürfen sich von technischen Problemen nicht beirren lassen und müssen gekonnt damit umgehen.

Abschließend bleibt zu sagen, dass der Workflow für jede Sendung aufgrund der Einteilung in Vorbereitung, Live-Sendung und Nachbearbeitung zwar immer sehr ähnlich wirken mag, jedoch müssen die Respeaker\*innen eine gewisse Flexibilität und hohe Konzentration an den Tag legen, um mit Änderungen im Sendungsablauf und möglichen Problemfaktoren zurechtzukommen.

## Conclusio

Ziel der vorliegenden Masterarbeit war es, herauszufinden, welche Arten von Sendungen von Live-Untertitler\*innen bearbeitet werden und wie der Workflow dabei aussieht. Zudem sollten damit in Verbindung stehende Problemfaktoren ausfindig gemacht werden.

Im ersten Teil der Arbeit wurden die theoretischen Grundlagen präsentiert. Dafür wurden die Themen audiovisuelle Translation, Untertitelung, Live-Untertitelung und Respeaking beleuchtet. Es wurde herausgearbeitet, dass es sich beim Respeaking um eine intralinguale Dolmetschleistung handelt. Nachfolgend wurden problemauslösende Faktoren und Bewältigungsstrategien beschrieben.

Im empirischen Teil wurde zuerst aufgezeigt, welche Sendungen von Respeaker\*innen Untertitelt werden. Anhand eines selbst entworfenen Rasters wurden die Sendungen einer Untertitelungsfirma in Programmsparte und Sendungsformat kategorisiert.

Die Analyse ergab, dass die Live-Untertitler\*innen vor allem kurze Programme (Dauer unter einer Stunde) bearbeiten. Die meisten Sendungen mit Live-Untertiteln wurden live ausgestrahlt, wobei darunter auch voraufgezeichnete Sendungen fielen. Für diese blieb jedoch nicht ausreichend Zeit, die Untertitel im Vorhinein anzufertigen.

Zudem hat sich herausgestellt, dass vorproduzierte und semi-live Untertitel am häufigsten gesendet wurden. Fast ein Drittel aller Sendungen wurde mit gemischten Untertiteln gezeigt, wohingegen nur wenige Sendungen komplett live Untertitelt wurden. Dies lässt den Schluss zu, dass fast immer Informationsmaterial von den Sendeanstalten zur Verfügung gestellt wurde und dieses sinnvoll für die Vorbereitung der Untertitel herangezogen werden konnte.

Es wurde außerdem herausgefunden, dass die Respeaker\*innen im überwiegenden Ausmaß Informationssendungen bearbeiten. Dazu zählten vor allem Magazinsendungen und Nachrichtenprogramme, die sich mit tagesaktuellen Ereignissen beschäftigten. Dass diese Programmsparte in der Untersuchung überproportional stark vertreten war, dürfte sich aber auch damit begründen lassen, dass Informationssendungen insbesondere bei öffentlich-rechtlichen Fernsehanstalten zu finden sind (vgl. Maurer et al. 2021), welche aufgrund der gesetzlichen Verpflichtungen dazu angehalten sind, eine gewisse Untertitelquote zu erfüllen. Im privaten Bereich werden eher Unterhaltungssendungen gezeigt (vgl. Maurer et al. 2021), wobei darunter die Sparte der fiktionalen Unterhaltung für die Live-Untertitelung ohnehin komplett wegfällt.

Die gemischte Untertitelungsmethode kam vor allem bei Magazinsendungen zum Einsatz. Dies lässt sich damit erklären, dass in solchen Programmen neben voraufgezeichneten Videos oft (Live-)Gespräche mit Gästen geführt werden, welche dementsprechend mittels Respeaking untitelt werden. Semi-live Untertitel waren besonders bei Nachrichtensendungen inklusive Wetterberichten zu finden. Reine live-produzierte Untertitel sind dahingegen eher im nonfiktionalen Unterhaltungsbereich oder bei Sportsendungen anzutreffen.

Der zweite Teil der Untersuchung beschäftigte sich mit dem Workflow für das Erstellen von Live-Untertiteln. Der Arbeitsablauf wurde ausgehend von dem Prozessmodell von Pöchhacker/ Remael (2019) in Vorbereitung, Live-Teil und Nachbearbeitung gegliedert. Mithilfe eines Beobachtungsrasters zur Selbstreflexion wurden die einzelnen Vorgänge beim Untertiteln einer konkreten Live-Sendung dokumentiert. Die Erkenntnisse wurden mit den Ergebnissen eines Expert\*inneninterviews mit zwei Respeaker\*innen abgeglichen und ergänzt.

Die Analyse lässt den Schluss zu, dass die Vorbereitung und vorhandenes Material im Zuge der Bearbeitung einer Sendung für die Live-Untertitelung essenziell sind. Je vollständiger und korrekter die Texte, umso mehr erleichtert es die Arbeit der Respeaker\*innen. Es ist davon auszugehen, dass dadurch auch die Qualität der Untertitel allgemein steigt, welche vorbereitet weniger fehleranfällig und wesentlich synchroner sind, als es live-produzierte Untertitel wären.

Darüber hinaus ist festzuhalten, dass die Respeaker\*innen für ihren Beruf nicht nur Fertigkeiten im simultanen (intralingualen) Dolmetschen benötigen, sondern auch technische Kompetenzen aufweisen sollten. Diese Komponente ist aus der Live-Untertitelung nicht wegzudenken, weil verschiedene Programme (teilweise gleichzeitig) benutzt werden müssen. Die entsprechenden Einstellungen sollten die Respeaker\*innen bestenfalls verinnerlicht haben, um den Arbeitsablauf zu beschleunigen und um möglichst wenig kognitive Ressourcen für die parallel ablaufenden Prozesse während des Live-Teils an technischer Stelle aufwenden zu müssen. In der Ausbildung sollten daher neben dem Erlangen von Dolmetsch-Fähigkeiten bzw. dem Fokus auf dem Multitasking und dem gleichzeitigen Verstehen des Ausgangstextes sowie Erstellen des Zieltextes ebenso technische Aspekte vermehrt miteinbezogen werden.

In einem weiteren Schritt wurden die von den Respeaker\*innen genannten Problemfaktoren während des Workflows festgehalten. Die Interviewstudie hat ergeben, dass besonders technische Herausforderungen die Arbeit der Live-Untertitler\*innen in allen Phasen erschweren. Dies hängt hauptsächlich mit den verwendeten Programmen für das Erstellen der Live-Untertitel oder mit Problemen bei Bild und Ton zusammen. Auch Problemfaktoren, die den *problem triggers* nach Gile (2009) entsprechen, wurden aufgezeigt. Darunter fallen etwa

schnelle Sprechgeschwindigkeit, Dialekte und Wortspiele. Die gewonnenen Erkenntnisse decken sich mit den Studienergebnissen von Szarkowska et al. (2019), Jorda (2021) und Stehle (2021). Es wäre deswegen sinnvoll, die problemauslösenden Faktoren für das Respeaking um ein technisches Element zu erweitern und auch den Ausgangstext bei der Fehlerbeurteilung für die Qualitätsmessung im NER-Modell zu berücksichtigen.

Generell lässt sich festhalten, dass Schwierigkeiten unterschiedlichster Art beim Workflow auftreten können. Die Respeaker\*innen lassen ihre Erfahrungen miteinfließen und wenden adäquate Bewältigungsstrategien an, um die Problemfaktoren zu umgehen. Es kann angenommen werden, dass sich die Problemfaktoren vervielfachen, sobald weniger Vorbereitungsmaterial vorhanden ist, worunter in weiterer Folge auch die Qualität leidet.

Inwieweit sich die Live-Untertitelung in Zukunft aufgrund des Auftretens der automatischen Spracherkennung weiterentwickelt, bleibt abzuwarten. Der Workflow mit den Phasen der Vorbereitung, Live-Sendung und Nachbearbeitung wird zwar weitestgehend erhalten bleiben, allerdings dürften sich die einzelnen Arbeitsschritte der Live-Untertitler\*innen abwandeln. Es ist jedoch davon auszugehen, dass die technische Komponente weiter zunimmt und Fertigkeiten diesbezüglich noch mehr an Relevanz gewinnen, weil sich die Problemfaktoren hinsichtlich der Technik mit der Weiterentwicklung nicht in Luft auflösen werden, sondern wohl eher noch häufiger anzutreffen sind.



## Bibliografie

- AIIC (2022). Dolmetscharten. <https://aiic.de/service/beratung/dolmetscharten> (Stand: 08.02.2022).
- Arumí Ribas, Marta/ Romero-Fresco, Pablo (2008). A practical proposal for the training of respeakers. *JoSTrans: Journal for Specialised Translation* 10, 106-127.
- Chmiel, Agnieszka/ Lijewska, Agnieszka/ Szarkowska, Agnieszka/ Dutka, Łukasz (2018). Paraphrasing in respeaking – comparing linguistic competence of interpreters, translators and bilinguals. *Perspectives* 26 (5), 725-744.
- Díaz Cintas, Jorge/ Matamala, Anna/ Neves, Josélia (2010). Media for all: New developments. In: Díaz Cintas et al. (Hg.), 11-22.
- Díaz Cintas, Jorge/ Matamala, Anna/ Neves, Josélia (Hg.) (2010). *New insights into audiovisual translation and media accessibility. Media for all 2*. Amsterdam/ New York: Rodopi.
- Díaz Cintas, Jorge/ Remael, Aline (2014). *Audiovisual translation: Subtitling*. London: Routledge.
- DSB (Deutsche Schwerhörigenbund e.V. DSB) (2021). Wir über uns. Aufgaben und Ziele. <https://www.schwerhoerigen-netz.de/informationen/wir-ueber-uns/wir-ueber-uns> (Stand 08.02.2022).
- EBU (European Broadcasting Union) (2004). EBU report on access services, EBU Technical-Information I144-2004. <https://tech.ebu.ch/docs/i/i044.pdf> (Stand: 03.02.2022).
- EU (2018). *Richtlinie 2018/1808 des Europäischen Parlaments und des Rates* vom 14. November 2018 zur Änderung der Richtlinie 2010/13/EU zur Koordinierung bestimmter Rechts- und Verwaltungsvorschriften der Mitgliedstaaten über die Bereitstellung audiovisueller Mediendienste (Richtlinie über audiovisuelle Mediendienste) im Hinblick auf sich verändernde Marktgegebenheiten. <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:32018L1808> (Stand: 08.02.2022).
- Eugeni, Carlo (2009). Respeaking the BBC news. A strategic analysis of respeaking on the BBC. *The Sign Language Translator and Interpreter* 3 (1), 29-68.

- Eugeni, Carlo (2010). Teaching respeaking to conference interpreters. *Intersteno Education Committee Archive*, 1-14. <https://www.intersteno.it/materiale/ComitScientifico/EducationCommittee/Russello2010Teaching%20Respeaking%20to%20Conference%20Interpreters.pdf> (Stand: 27.01.2022).
- Geens, Rita (2007). Live subtitling through respeaking: A new discipline in interpretation? University College Ghent. <https://pdffox.com/live-subtitling-through-respeaking-pdf-free.html> (Stand: 22.05.2022).
- Gehrau, Volker (2001). *Fernsehgenres und Fernsehgattungen: Ansätze und Daten zur Rezeption, Klassifikation und Bezeichnung von Fernsehprogrammen*. München: Reinhard Fischer.
- Gerzymisch-Arbogast, Heidrun (2013). Gutachten zu der Bezeichnung Schriftdolmetschen. [https://bsd-ev.org/wp-content/uploads/2019/09/gutachten\\_gerzymisch.pdf](https://bsd-ev.org/wp-content/uploads/2019/09/gutachten_gerzymisch.pdf) (Stand: 16.01.2022).
- Gile, Daniel (2009). *Basic concepts and models for interpreters and translator training*. Amsterdam/ Philadelphia: John Benjamins.
- Gleason, Henry. A. (1965). *Linguistics and English grammar*. New York [u.a.]: Holt, Rinehart and Winston.
- Gottlieb, Henrik (2002). Untertitel: Das Visualisieren filmischen Dialogs. In: Friedrich, Hans-Edwin/ Jung, Uli (Hg.) *Schrift und Bild im Film*. Bielefeld: Aisthesis, 185-214.
- Helfferich, Cornelia (2014). Leitfaden- und Experteninterviews. In: Baur, Nina/ Blasius, Jörg (Hg.) *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer VS, 559-574.
- Hezel, Susanne (2009). Untertitelung für Hörgeschädigte für das deutsche Fernsehen: Vorgehensweisen, Forderungen, Richtlinien. In: Nagel et al. (Hg.), 147-264.
- ILSA (o.J.). About the project. <http://ka2-ilsa.webs.uvigo.es/project> (Stand: 22.05.2022).
- Ingram, Robert (1974/2015). A communication model of the interpreting process. In: Roy, C./ Napier, J. (Hg.) *The sign language interpreting studies reader*. Amsterdam/ Philadelphia: John Benjamins Publishing, 22-28.

- Ivarsson, Jan/ Carroll, Mary (1998). *Subtitling*. Simrishamn: TransEdit.
- Jakobson, Roman (1959). On linguistic aspects of translation. In: Venuti, Lawrence (Hg.) (2000) *The translation studies reader*. London/ New York: Routledge, 113-118.
- Jakobson, Roman (1971). *Fundamentals of language*. Den Haag: Mouton.
- Jekat, Susanne Johanna (2014). Respeaking: Syntaktische Aspekte des Transfers von gesprochener Sprache in geschriebene Sprache. In: Jekat, Susanne Johanna/ Jüngst, Heike Elisabeth/ Schubert, Klaus/ Villiger, Claudia (Hg.) *Sprache barrierefrei gestalten. Perspektiven aus der Angewandten Linguistik*. Berlin: Frank & Timme, 87-108.
- Jorda, Birgit (2021). *Intralinguales Respeaking für die Live-Untertitelung: Eine Fallstudie zu Problemauslösern und Strategien*. Masterarbeit, Universität Wien.
- Jüngst, Heike (2010). *Audiovisuelles Übersetzen. Ein Lehr- und Arbeitsbuch*. Tübingen: Narr.
- Käber, Marco (2016). Die Geschichte der Liveuntertitelung – ein Praxisbericht. In: Mälzer, Nathalie (Hg.), 295-305.
- Kade, Otto (1968). *Zufall und Gesetzmäßigkeit in der Übersetzung*. Beihefte zur Zeitschrift Fremdsprachen. 1. Leipzig: VEB Verlag Enzyklopädie.
- Kalina, Sylvia (1998). *Strategische Prozesse beim Dolmetschen*. Tübingen: Gunter Narr.
- Kalina, Sylvia (2005). Quality assurance for interpreting processes. *Meta* 50 (2), 768-784.
- Kirchhoff, Hella (1976/2002). Simultaneous interpreting: Interdependence of variables in the interpreting process, interpreting models and interpreting strategies. In: Pöchhacker, Franz/ Shlesinger, Miriam (Hg.) *The interpreting studies reader*. London/ New York: Routledge, 112-130.
- Lambourne, Andrew (2006). Subtitle respeaking. A new skill for a new age. In: Eugeni, Carlo/ Mack, Gabriele (Hg.) *Intralinea, Special Issue on Respeaking*. [http://www.intralinea.org/specials/article/Subtitle\\_respeaking](http://www.intralinea.org/specials/article/Subtitle_respeaking) (Stand: 22.05.2022).
- Lambourne, Andrew (2007). *Realtime subtitling: extreme audiovisual translation*. Paper presented at the conference LSP Translation Scenarios in Vienna in May 2007.

- Lambourne, Andrew/ Hewitt, Jill/ Lyon, Caroline/ Warren, Sandra (2004). Speech-based real-time subtitling services. *International Journal of Speech Technology* 7 (4), 269-279.
- Lamnek, Siegfried (2010). *Qualitative Sozialforschung*. Weinheim/Basel: Beltz.
- Lindner, Peter (2016). Vorproduktion – Semi-Live – Live. Ein Arbeitstag in der Untertitelredaktion des NDR. In: Mälzer, Nathalie (Hg.), 307-316.
- Luyckx, Bieke/ Delbeke, Tiejts/ Van Waes, Luuk/ Leijten, Mariëlle/Remael, Aline (2010). Live subtitling with speech recognition. Causes and consequences of text reduction. Working Papers, University of Antwerp, Faculty of Business and Economics.
- Maaß, Christiane/ Rink, Isabel (Hg.) (2020). *Handbuch Barrierefreie Kommunikation. Kommunikation – Partizipation – Inklusion*. Berlin: Frank & Timme.
- Mälzer, Nathalie (Hg.) (2016). *Barrierefreie Kommunikation – Perspektiven aus Theorie und Praxis*. Berlin: Frank & Timme.
- Mälzer, Nathalie/ Wünsche, Maria (2020). Untertitelung für Hörgeschädigte (SDH). In: Maaß, Christiane/ Rink, Isabel (Hg.), 327-344.
- Mankauskienė, Dalia (2018). *Problem triggers in simultaneous interpreting from English into Lithuanian*. Summary of Doctoral Dissertation, Vilnius University. [https://www.researchgate.net/publication/325319733\\_Problem\\_triggers\\_in\\_simultaneous\\_interpreting\\_from\\_English\\_into\\_Lithuanian\\_Summary\\_of\\_my\\_PhD\\_Thesis](https://www.researchgate.net/publication/325319733_Problem_triggers_in_simultaneous_interpreting_from_English_into_Lithuanian_Summary_of_my_PhD_Thesis) (Stand: 20.05.2022).
- Marsh, Alison (2004). *Simultaneous interpreting and respoken: A comparison*. Unpublished MA Thesis, University of Westminster.
- Martínez Pérez, Juan (2012). Applying a punctuation-based segmentation to a new add-on display mode of respoken subtitles. In: Remael, Aline/ Orero, Pilar/ Carroll, Mary (Hg.) *Audiovisual translation and media accessibility at the crossroads. Media for all 3*. Amsterdam/ New York.: Rodopi, 132-146.
- Maurer, Torsten/ Wagner, Matthias/ Weiß, Hans-Jürgen (2021). Programmprofile von Das Erste, ZDF, RTL, VOX, Sat.1 und ProSieben. Ergebnisse der ARD/ZDF-Programmanalyse 2020 – Teil 1. *Media Perspektiven* 4/2021, 240-266.

- Nagel, Silke (2009). Das Übersetzen von Untertiteln: Prozess und Probleme der Kurzfilme Shooting Bokkie, Wasp und Green Bush. In: Nagel et al. (Hg.), 23-146.
- Nagel, Silke/ Hezel, Susanne/ Hinderer, Katharina/ Pieper, Katrin (Hg.) (2009). *Audiovisuelle Übersetzung: Filmuntertitelung in Deutschland, Portugal und Tschechien*. Frankfurt am Main [u.a.]: Lang.
- Neves, Josélia (2005). *Audiovisual translation: subtitling for the deaf and hard-of-hearing*. Dissertation, University of Surrey-Roehampton.
- Neves, Josélia (2009). Interlingual subtitling for the deaf and hard-of-hearing. In: Díaz Cintas, Jorge/ Anderman, Gunilla (Hg.) *Audiovisual translation: Language transfer on screen*. Basingstoke: Palgrave Macmillan, 151-169.
- Orero, Pilar (2006). Real-time subtitling in Spain. In: Eugeni, Carlo/ Mack, Gabriele (Hg.) *Intralinea, Special Issue on Respeaking*. <http://www.intralinea.org/specials/article/1689> (Stand: 03.02.2022).
- Platter, Judith (2015). *Translation im Spannungsbereich von Mündlichkeit und Schriftlichkeit, Schriftdolmetschen in Österreich. Eine textbasierte Analyse*. Dissertation, Universität Wien.
- Pöchhacker, Franz/ Remael, Aline (2019). New efforts?: A competence-oriented task analysis of interlingual live subtitling. *Linguistica Antverpiensia, New Series: Themes in Translation Studies* 18, 130-143.
- Prillwitz, Siegmund (2001). *Angebote für Gehörlose im Fernsehen und ihre Rezeption*. Kiel: Unabhängige Landesanstalt für das Rundfunkwesen (URL).
- Remael, Aline (2012). Media accessibility. In: Gambier, Yves/ Doorslaer, Luc van. *Handbook of translation studies: Volume 3*. Amsterdam/ Philadelphia: John Benjamins, 95-101.
- Remael, Aline/ Van Waes, Luuk/ Leitjen, Marielle (2014). Live subtitling with speech recognition: How to pinpoint the challenges. In: Abend-David, Dror (Hg.) *Media and translation. An interdisciplinary approach*. New York/ London: Bloomsbury, 121-148.
- Robson, Gary D. (2004). *The closed captioning handbook*. Amsterdam: Elsevier.

- Romero-Fresco (2015). *The reception of subtitles for the deaf and hard of hearing in Europe*. Bern: Peter Lang.
- Romero-Fresco, Pablo (2009). More haste less speed: Edited versus verbatim respoken subtitles. *Vial-Virgo International Journal of Applied Linguistics* 6, 109-133.
- Romero-Fresco, Pablo (2010). Standing on quicksand: hearing viewers' comprehension and reading patterns of respoken subtitles for the news. In: Díaz Cintas et al. (Hg.), 175-194.
- Romero-Fresco, Pablo (2011). *Subtitling through speech recognition: Respeaking*. Manchester/Kinderhook: St. Jerome.
- Romero-Fresco, Pablo (2019). Respeaking: subtitling through speech recognition. In: Pérez-González, Luis (Hg.) *The Routledge handbook of audiovisual translation*. Abingdon/New York: Routledge, 96-113.
- Romero-Fresco, Pablo/ Martínez Pérez, Juan (2015). Accuracy rate in live subtitling: The NER model. In: Baños Piñero, Rocío/ Díaz Cintas, Jorge (Hg.) *Audiovisual translation in a global context: Mapping an ever-changing landscape*. London: Palgrave Macmillan, 28-50.
- Romero-Fresco, Pablo/ Pöchhacker, Franz (2017). Quality assessment in interlingual live subtitling: The NTR model. *Linguistica Antverpiensia New Series: Themes in Translation Studies* 14, 149-167.
- Stehle, Johannes (2021) *Anwendung des NER-Modells für deutschsprachige Live-Untertitel*. Masterarbeit, Universität Wien.
- Szarkowska, Agnieszka/ Dutka, Łukasz/ Pilipczuk, Olga/ Krejtz, Krzysztof (2019). Respeaking crisis points. An exploratory study into critical moments in the respeaking process. In: Deckert, Mikołaj (Hg.) *Audiovisual translation: Research and use*. Berlin: Peter Lang, 193-215.
- Szarkowska, Agnieszka/ Krejtz, Krzysztof/ Dutka, Łukasz/ Pilipczuk, Olga (2018). Are interpreters better respeakers? *The Interpreter and Translation Trainer* 12 (2), 207-226.
- Tiittula, Liisa (2006). Schriftdolmetschen – Mündlichkeit im schriftlichen Gewand. In: Suominen, M./ Arppe, A./ Airola, A./ Heinämäki, O./ Miestamo, M./ Määttä, U./ Niemi, J./

Pitkänen, K. K./ Sinnemäki, K. (Hg.) *A man of measure. Festschrift in honour of Fred Karlsson on his 60th birthday. SKY journal of linguistics* 19, 481-488.

UN-BRK (2008). Übereinkommen der Vereinten Nationen über die Rechte von Menschen mit Behinderung. [https://www.behindertenbeauftragte.de/SharedDocs/Publikationen/UN\\_Konvention\\_deutsch.pdf?\\_\\_blob=publicationFile&v=2](https://www.behindertenbeauftragte.de/SharedDocs/Publikationen/UN_Konvention_deutsch.pdf?__blob=publicationFile&v=2) (Stand: 03.02.2022).

Witzel, Jutta (2020). Ausprägungen und Dolmetschstrategien beim Schriftdolmetschen. In: Maaß, Christiane/ Rink, Isabel (Hg.), 303-325.

# Anhang

## Interviewleitfaden

### Thema 1: Vorbereitung

- Wie sah die Vorbereitung für die Untertitelung der Live-Sendung bei Ihnen aus?
- Wie viele Wörter haben Sie vorbereitet und schlussendlich gesprochen?
- Welche Art von Wörtern haben Sie gesprochen?
- Wie schätzen Sie ab, wie viel Vorbereitungszeit notwendig ist?
- Gibt es etwas, was Ihre Vorbereitung erschwert?
- (Was sind die Gründe für die auftretenden Probleme?)
- (Wie häufig sind Sie damit konfrontiert?)

### Thema 2: Live-Sendung

- Beschreiben Sie bitte, was Sie während der Live-Sendung gemacht haben.
- Waren Beitragstexte und Anmoderationen, die vorbereitet werden konnten, brauchbar/vollständig?
- Gibt es etwas, was Ihre Arbeit während der Live-Sendung erschwert?
- (Was sind die Gründe für die auftretenden Probleme?)
- (Wie häufig sind Sie damit konfrontiert?)

### Thema 3: Nachbearbeitung

- Was musste bei der live-untertitelten Sendung nachbearbeitet werden?
- Wofür dient die Nachbearbeitung in dem Fall?
- Gibt es etwas, was die Nachbearbeitung erschwert?
- (Was sind die Gründe für die auftretenden Probleme?)
- (Wie häufig sind Sie damit konfrontiert?)

Haben Sie sonst noch Anmerkungen zum Workflow beim Live-Untertiteln?



## Transkription der Interviews

### 1 Interviewperson 1 (IP1)

2 *F: Ich schreibe über den Workflow bei der Live-Untertitelung und habe ihn in die Bereiche*  
3 *Vorbereitung, Live-Sendung und Nachbearbeitung aufgegliedert. Dazu werde ich Fragen stel-*  
4 *len und beginne gleich mit der ersten: Wie sah die Vorbereitung für die Untertitelung der Live-*  
5 *Sendung, die Sie heute gemacht haben, bei Ihnen aus?*

6 A: Es gab teilweise Texte, die ich verwendet habe, um einzelne Beiträge vorzubereiten. Es gab  
7 aber auch einen Teil, wo nur ein Video verfügbar war, wo ich selbst den Text mittels Sprach-  
8 erkennung hinzugefügt habe. Für den Live-Teil, wobei ich nicht wusste, ob der wirklich wäh-  
9 rend der Sendung live war zu dem Zeitpunkt, habe ich elf Vokabeln mir herausgeschrieben, um  
10 diese noch der Spracherkennungssoftware beizubringen. Das Video kam vorab, das heißt, ich  
11 habe diesen Teil nicht live-live gemacht, sondern im Endeffekt vorbereiten können. Das habe  
12 ich auch mit Spracherkennung gemacht. Die Vorbereitungsdauer war für die Sendung, die eine  
13 Dreiviertelstunde gedauert hat, ca. zwei Stunden, zwei bis drei Stunden lang. Ich denke, das ist  
14 alles, was ich zur Vorbereitung sagen kann.

15 *F: Wie haben Sie entschieden, welche Wörter Sie einsprechen und welche Art von Wörtern war*  
16 *das?*

17 A: Ich kenne mein Spracherkennungsprogramm schon recht gut und daher weiß ich, dass man-  
18 che Dinge sicherlich gut erkannt werden, und habe mich auf Eigennamen fokussiert, sowohl  
19 mit Titel bzw. mit Anrede, also „Frau etc.“ oder „Mann etc.“ Einmal mit Vor- und Nachnamen.  
20 Dann kam auch ein Hund vor. Dessen Namen habe ich auch ausgesprochen, und ein paar geo-  
21 graphische Begriffe waren dabei, wo ich nicht sicher war, ob es gut erkannt werden würde, und  
22 zwei medizinische Begriffe ebenfalls. In dem Fall waren es also nur Nomen, die ich ausgespro-  
23 chen habe.

24 *F: Sie haben erwähnt, wie viel Zeit Sie in die Vorbereitung gesteckt haben. Wie schätzen Sie*  
25 *ab, wie viel Vorbereitungszeit notwendig ist?*

26 A: Das ist sehr abhängig von der Sendung, grundsätzlich aber ist es davon abhängig, ob Spre-  
27 chertexte zur Verfügung stehen, wie gut diese sind, ob sie auch schon O-Töne inkludieren oder  
28 nur den Sprechertext. Davon mache ich das abhängig, und die Erfahrung zeigt dann, wie lange  
29 man ungefähr brauchen kann. Ebenso ist relevant, ob Videomaterial zur Verfügung steht, mit  
30 dem man das vorab abgleichen kann.

31 *F: Gleichen Sie das Material ab mit den Videos, wenn es zur Verfügung steht?*

32 *A: Ja, außer es bleibt keine Zeit.*

33 *F: In welcher Reihenfolge arbeiten Sie die verfügbaren Texte ab?*

34 *A: In der Reihenfolge der Vollständigkeit. Also wenn ich einen Beitrag habe, wo ein*  
35 *Sprechertext und ein Video verfügbar ist, dann bearbeite ich diesen zuerst. Wenn dann kein*  
36 *Video mehr verfügbar ist, würde ich zuerst einen Text bearbeiten, der einen Sprechertext hat,*  
37 *und gegen Ende erst einen Teil, wo ich nur ein Video zur Verfügung habe.*

38 *F: Gibt es etwas, was die Vorbereitung erschwert und wenn ja, was sind die Gründe dafür?*

39 *A: Andere Sendungen, mit denen ich beauftragt war. In vielen Fällen arbeitet man nicht nur an*  
40 *einer Sendung, sondern an mehreren gleichzeitig. Das verändert natürlich die Art der Vorbe-*  
41 *reitung oder man zieht manche Beiträge vielleicht vor. Sonst denke ich nicht, dass es Einschrän-*  
42 *kungen gab. Manchmal weiß man in einem Sprechertext nicht, ob ein O-Ton nur von einer*  
43 *Person gesprochen wird. Das hat eventuell eine Auswirkung auf die Einfärbung, wenn man mit*  
44 *Untertiteln arbeitet, die bei O-Tönen eingefärbt werden.*

45 *F: Dann kommen wir zur Live-Sendung. Beschreiben Sie bitte, was Sie während der Live-Sen-*  
46 *dung gemacht haben.*

47 *A: Der Start der Live-Sendung wird immer markiert, damit der Zuseher, die Zuseherin auch*  
48 *weiß, dass jetzt eine Live-Untertitelung folgt. Dann hört man eine Millisekunde mit, genau die*  
49 *Sekundenanzahl, die ich jetzt höre, um festzustellen, ob der vorbereitete Text passt. Das kann*  
50 *ich jetzt nicht nennen. Und im Endeffekt schicke ich mit Enter meine vorbereiteten Untertitel*  
51 *raus, warte dann, dass die Standzeit auch stimmt. Dabei zeigt mir mein Untertitelprogramm, ob*  
52 *der Untertitel lang genug steht, indem ein Balken sich in einen grünen Bereich bewegt. Dann*  
53 *muss man von Beitrag zu Beitrag im Programm springen, entweder durch Enter – beim letzten*  
54 *Untertitel kommt man in den nächsten Ordner – oder man kann auch durch Klicken in den*  
55 *nächsten Ordner kommen.*

56 *Wenn ein Fehler im Untertitel, der vorbereitet war, aufkommt, versucht man diesen noch*  
57 *schnell zu korrigieren, bevor man ihn mit Enter rausschickt. Teilweise habe ich auch während*  
58 *der Sendung noch meine O-Töne eingefärbt, weil ich dazu nicht genug Zeit hatte oder nicht*  
59 *wusste, dass eine Person interviewt werden würde. Bei dieser spezifischen Sendung heute gab*  
60 *es einen O-Ton, der nicht markiert war in einem Beitrag. Das waren ungefähr drei Sätze, wo*  
61 *ich dann umschalten musste auf ein „Sprachinterface“, wo ich dann direkt einsprechen kann*

62 und meinen Untertitel hinausschicken kann. Dann schalte ich wieder um auf den Modus der  
63 vorbereiteten Untertitel und schicke weiter meine Untertitel raus. Und am Ende der Sendung  
64 gibt es eine Endmarkierung für das Ende der Untertitel.

65 *F: Waren die Beitragstexte und An- bzw. Abmoderation brauchbar bzw. vollständig?*

66 A: Ja, in diesem Fall waren die Anmoderationen für die Beiträge vollständig. Ein paar Füllworte  
67 wurden in der Live-Situation hinzugefügt. Die habe ich aber nicht beachtet, weil sie für den  
68 Inhalt nicht relevant waren.

69 *F: Gibt es etwas, was die Arbeit während der Live-Sendung für Sie erschwert?*

70 A: Technikprobleme, sowohl Ton als auch Bild können ausfallen, das Programm kann zu lang-  
71 sam sein, sowohl das Untertitelprogramm als auch das Spracherkennungsprogramm. Das sind  
72 die größten Probleme, die auftreten. Gegebenenfalls wird bei Sendungen manchmal die Rei-  
73 henfolge von Beiträgen verändert, was auch zu Verwirrung führen kann.

74 *F: Wie häufig sind Sie mit solchen Problemen konfrontiert bei einer Live-Sendung?*

75 A: Nicht notwendigerweise bei jeder Live-Sendung, aber jeden zweiten Tag.

76 *F: Dann kommen wir zur Nachbearbeitung. Was musste bei der live-untertitelten Sendung  
77 nachbearbeitet werden?*

78 A: Vorher habe ich noch vergessen, dass ich während der Live-Sendung auch meist aufschreibe,  
79 wenn mir schon Fehler auffallen, die leider dennoch auf Sendung gegangen sind. Das heißt,  
80 hier schreibe ich mir immer eine kurze Wortfolge auf, damit ich dann mittels Suchfunktion  
81 schneller an die Stellen komme, die nachbearbeitet werden müssen vom Textinhalt. Diese ma-  
82 che ich oft zuerst. Die bessere ich zuerst aus. Dann setze ich den ersten Untertitel mit dem  
83 richtigen Timecode und kontrolliere nochmal alle Timecodes, alle Untertitel auf die Standzeit,  
84 ob die auch korrekt ist und ob der Untertitel an der richtigen Stelle erscheint.

85 *F: Wofür dient die Nachbearbeitung in dem Fall?*

86 A: Für ein Archiv, das nie benützt wird. Und für eine weitere Ausstrahlung, die noch folgt in  
87 der Nacht.

88 *F: Gibt es etwas, was die Nachbearbeitung erschwert?*

89 A: Wenn ein Mitschnitt nicht verfügbar ist oder unvollständig ist oder spät kommt. Oder auch,  
90 wenn eine die Untertiteldatei selbst einen Fehler aufweist.

91 *F: Was sind die Gründe für die genannten Probleme?*

92 A: Technische Probleme, entweder beim Untertitel-Dienstleister oder bei der Sendeanstalt.

93 *F: Wie häufig sind Sie mit solchen Problemen konfrontiert?*

94 A: Jeden zweiten Tag.

95 *F: Haben Sie sonst noch Anmerkungen zum Workflow beim Live-Untertiteln?*

96 A: Ich würde sagen, sowohl bei der Vorbereitung als auch bei der Nachbearbeitung und wäh-  
97 rend der Live-Sendung gibt es viele Faktoren, die beachtet werden müssen. Die sprachliche  
98 Natur ist eventuell das geringste Übel.

99 *F: Inwiefern?*

100 A: Dass es einfach sehr viele technische Dinge gibt, die passieren können, die nicht wirklich  
101 mit der Sprachdienstleistung selbst zu tun haben.

102 *F: Welche technische Probleme kommen am häufigsten vor bzw. erschweren die Arbeit am*  
103 *meisten?*

104 A: Hauptsächlich Fehlen von Bild und Ton oder auch Ausfälle, die den Computer betreffen und  
105 gar nicht die Sendeanstalt, oder die anderen Hardwares, mit denen wir arbeiten.

106 *F: Sonst noch Anmerkungen?*

107 A: Nein.

108 **Interviewperson 2 (IP2)**

109 *F: Vielen Dank, dass Sie sich Zeit genommen haben für das Interview. Es geht um den Workflow*  
110 *beim Live-Untertiteln. Ich habe ihn in drei Bereiche gegliedert, Vorbereitung, Live-Sendung*  
111 *und Nachbearbeitung, und stelle dazu Fragen. Thema 1 ist die Vorbereitung. Wie sah die Vor-*  
112 *bereitung für die Untertitelung der Live-Sendung bei Ihnen heute aus?*

113 A: Ich bin ins Büro gekommen und habe erst einmal die Sendung mit dem System verknüpft,  
114 damit ich alle Texte und Videos habe, damit ich sehen kann, was schon vorbereitet ist von der  
115 Redaktion und was noch fehlt, ob irgendwelche spezifischen Eigennamen vorkommen und was  
116 das grundsätzliche Thema der Sendung ist. Dann habe ich mein Sprachprofil für die Spracher-  
117 kennung kopiert. Das dauert nämlich immer einige Zeit, weil es schon recht groß ist. Deswegen  
118 versuche ich das immer ganz am Anfang zu machen, damit ich – während alle anderen Systeme  
119 laden – das schon kopiere und es bereitsteht, spätestens wenn die Sendung beginnt. Dann habe  
120 ich mir eben im System angesehen, welche Beiträge dieser Sendung schon fertig sind und diese  
121 zuerst vorbereitet. Meistens schaue ich einfach, wo das Video verfügbar ist, und kontrolliere  
122 dann die Texte, die mir zur Verfügung stehen seitens der Redaktion gleich mit dem Video,  
123 während ich das Ganze vorbereite, damit etwaige O-Töne – die stellen meistens das größte  
124 Problem dar – dann auch schon stimmen in der Live-Situation. So bin ich dann die einzelnen  
125 Beiträge, An- und Abmoderationen durchgegangen.

126 Heute in der Sendung war ein Kabarettist zu Gast. Dann habe ich mir auch den Namen seines  
127 aktuellen Programms bzw. seiner neuen Sendung eingespeichert, habe aber offensichtlich nicht  
128 genau genug gelesen. Denn der Titel ist ein Wortspiel und ich habe das eigentliche Wort ge-  
129 nommen, und nicht so, wie die Sendung heißt, nämlich das Wortspiel. Das musste ich dann in  
130 der Live-Situation ausbessern. Da hätte ich mich gleich besser vorbereiten sollen und genauer  
131 lesen sollen. Es war so circa zehn bis 20 Minuten Live-Teil, wo das Interview live im Fernsehen  
132 ausgestrahlt wurde. Und ich hatte aber nur an die 20 bis 30 Wörter, weil vieles sehr allgemein  
133 gehalten wurde. Es wurde auf die aktuelle politische Lage und gesundheitliche Lage auf der  
134 Welt eingegangen. Dazu habe ich die Wörter schon, weil die im Fernsehen natürlich ständig  
135 vorkommen.

136 *F: Welche Wörter haben Sie vorbereitet und wie viele haben Sie schlussendlich eingesprochen?*  
137 *Welche Art von Wörtern war das?*

138 A: Das waren unterschiedliche Wörter, wie Nomen und Verben, bei denen ich davon ausgehe,  
139 dass die Spracherkennung sie nicht kennt. So etwas wie „Atomkrieg“, zum Beispiel, habe ich

140 in meiner Liste gesehen, aber auch Eigennamen von einer Band oder der Name des Gastes am  
141 nächsten Tag, natürlich auch der Name des Gastes selber und seiner Rolle, die er spielt in sei-  
142 nem Kabarettprogramm. Hauptsächlich handelt es sich um Eigennamen, aber auch alle Verben  
143 und Nomen, die nicht dem Standardrepertoire der Spracherkennung angehören, wobei das im  
144 Prinzip in meinem eigenen Ermessen liegt, wie ich die einschätze, was ich glaube, was die  
145 Spracherkennung nicht kann. Wenn ich mich recht erinnere, wurden dann auch nicht alle ein-  
146 gesprochen. Das heißt, die Spracherkennung kannte ein paar von diesen Wörtern schon. Aber  
147 welche genau, das weiß ich dann nicht. Darauf achte ich nicht. Aber es waren weniger, als ich  
148 herausgeschrieben habe.

149 *F: Wie wissen Sie, wie viel Vorbereitungszeit notwendig ist?*

150 A: Das ist eigentlich sehr abhängig davon, wie lang eine Sendung ist, wie die Sendung aufge-  
151 baut ist. Also ob es viele vorbereitete Videos gibt und nur ein bisschen Moderation dazwischen.  
152 Oder ob die ganze Sendung live ist. Die heutige Sendung war 45 Minuten und davon eben zehn  
153 bis 20 Minuten richtig live. Und da war meine Vorbereitung von 16:30 Uhr bis 18:30 Uhr im  
154 Prinzip. Das war aber dem geschuldet, dass die Texte heute zum Beispiel nicht sehr gut waren.  
155 Es geht auch schneller, wenn die Texte, die ich von der Redaktion bekomme, besser sind. An-  
156 dererseits gibt es Sendungen, die komplett live sind, und da habe ich oft wenig Information, die  
157 ich ausschreiben kann. Und dann ist die Vorbereitung eigentlich kürzer, weil ich nur Wörter  
158 ausschreibe und einspreche und keine Texte vorbereiten muss.

159 *F: Gibt es etwas, was Ihre Vorbereitung erschwert? Was sind die Gründe dafür?*

160 A: Am erswerlichsten sind fehlerhafte und mangelnde Texte, wenn viele Passagen fehlen,  
161 was meistens O-Töne sind oder wenn diese falsch transkribiert werden, oder eben komplett  
162 fehlerhafte Texte.

163 *F: Wie häufig sind Sie mit solchen Problemen konfrontiert?*

164 A: Bei jeder Sendung. Es ist sehr selten, dass es kaum störende Elemente gibt. Meine Vermu-  
165 tung ist, dass seitens des Redaktionsteams nicht viel Wert darauf gelegt wird, ob die Texte  
166 grammatikalisch korrekt sind und die Zeichensetzung korrekt ist.

167 *F: Dann kommen wir zur Live-Sendung. Beschreiben Sie bitte, was Sie während der Live-Sen-  
168 dung gemacht haben.*

169 A: Der Anfang der Sendung beginnt immer mit einer Moderation live, aber im Prinzip habe ich  
170 alle Texte bis kurz vorm Schluss vorbereitet in zweizeilige Untertitel. Die sende ich dann mit-  
171 tels Enter-Taste live hinaus und kann auf meinem Rückbild kontrollieren, ob diese auch wirk-  
172 lich auf Sendung gehen. Ich muss, wenn es einen O-Ton gibt, diesen einfärben, falls ich das  
173 nicht schon vorher in der Vorbereitung gemacht habe. Und ich muss bei Inserts, bei Namen  
174 oder Reporter-Inserts, die Untertitel hochstellen, damit man das dann im Fernsehen trotzdem  
175 noch alles lesen kann. Im Idealfall kontrolliere ich auch live noch, ob Fehler drinnen sind, was  
176 sehr häufig auch noch der Fall ist, weil ich irgendetwas übersehen habe. Und dann bessere ich  
177 das schnell aus, bevor ich es wegschicke. Und dann gegen Ende der Sendung – das war bei  
178 dieser spezifischen Sendung der Fall – ist eben das Live-Interview. Da wechsele ich dann in den  
179 Live-Modus, spreche das Gesprochene nach, kontrolliere das, muss auch Farben wählen für  
180 diverse Sprecher\*innen, muss auch hochstellen – aber das natürlich alles, während ich zuhöre,  
181 spreche, lese, abschicke.

182 *F: Gibt es etwas, was Ihre Arbeit während der Live-Sendung erschwert?*

183 A: Ja, das war auch heute der Fall, dass sich plötzlich meine Ansicht am PC verkleinert hat und  
184 ich alle offenen Programme am PC gesehen habe. Dadurch hat natürlich die Spracherkennung  
185 nicht in mein Untertitelprogramm gesprochen. Dann musste ich manuell wieder reinklicken  
186 und alles wiederholen. Oder dass einfach die Spracherkennung manchmal Teile des Satzes nicht  
187 erkennen will oder komplett falsch erkennt, und ich das auch im Idealfall ausbesser und nach-  
188 hole. Es ist auch immer schwierig, wenn wie heute ein Kabarettist zu Gast ist, und dann even-  
189 tuell irgendwelche Witze schnell erzählt. Denn oft kommt es da auf die Geschwindigkeit an,  
190 wenn ein Witz erzählt wird. Das ist live natürlich sehr schwierig. Und Dialekte können auch  
191 ein Problem darstellen.

192 *F: Was sind die Gründe für die auftretenden Probleme?*

193 A: Das sind Probleme einerseits technischer Natur, dass die Technik streikt und das Ganze  
194 erschwert, und andererseits Probleme vom Ausgangsmedium, also von der Sendung oder was  
195 auch immer, wie Tonprobleme oder diverse Probleme bei den Sprecher\*innen, für mich also  
196 Geschwindigkeit oder Sprachfärbung.

197 *F: Wie häufig sind Sie mit solchen Problemen konfrontiert?*

198 A: Auch das ist eigentlich fast in jeder Sendung anzutreffen. Das kommt natürlich auf die Sen-  
199 dung an. Im Regionalfernsehen kommt mehr Dialekt vor als im überregionalen Fernsehen. Da

200 ist die Sprachfärbung nicht so das Problem. Dafür gibt es im Entertainmentbereich, was über-  
201 regional ist, oft schnelle Sprache, Umgangssprache, Anglizismen. Aber im Endeffekt ist das  
202 eine oder andere Problem in jeder Sendung zu finden. Die Frage ist, ob man sich davon aus der  
203 Ruhe bringen lässt oder nicht.

204 *F: Dann kommen wir zur Nachbearbeitung. Was musste bei der live-untertitelten Sendung*  
205 *nachbearbeitet werden?*

206 A: Das ist abhängig von den Auftraggeber\*innen. Bei der Sendung heute habe ich schon wäh-  
207 renddessen einige Fehler ausbessern können, sodass ich mir die nicht aufschreiben musste.  
208 Sonst sind es oft Beistrichfehler oder große oder kleine Anfangsbuchstaben, die ich gerne noch  
209 ausbesser. Dann muss ich darauf achten, dass jeweils beim Wechsel vom voraufgezeichneten  
210 Video zum Moderator der erste Untertitel immer genau am Anfang der Tonspur ist. Und beim  
211 Live-Teil habe ich natürlich mehr Fehler ausbessern müssen, einige Sachen hinzugefügt oder  
212 korrigiert, wenn ich etwas nicht ganz richtig verstanden habe. Und oft verschiebe ich auch noch  
213 die Untertitel an den einen richtigen Timecode, damit nicht eine ganz so große Verzögerung da  
214 ist, und das später in etwaigen Videotheken oder live im Fernsehen einfacher zu lesen ist.

215 *F: Wofür dient die Nachbearbeitung in dem Fall?*

216 A: In diesem Fall gibt es eine Wiederholung am nächsten Tag in der Früh im Fernsehen. Da  
217 wird auf jeden Fall die nachbearbeitete Version verwendet, und vielleicht wird sie auch in der  
218 Mediathek verwendet. Aber das weiß ich nicht.

219 *F: Gibt es etwas, was die Nachbearbeitung erschwert?*

220 A: Wenn die Situation live nicht gut war oder eben technische Probleme entstanden sind, muss  
221 ich hinterher fehlende Passagen einfügen. Oder wenn die Technik sehr langsam ist. Aber bei  
222 der konkreten Sendung eigentlich nicht.

223 *F: Was sind die Gründe für die auftretenden Probleme bei der Nachbearbeitung?*

224 A: Meistens Probleme in der Live-Situation. Je besser vorbereitet ist und je besser es live läuft,  
225 desto unproblematischer läuft die Nachbearbeitung. Je ungenauer man davor gearbeitet hat,  
226 desto umständlicher und länger wird die Nachbearbeitung. Aber es kommt auch darauf an, wo-  
227 für diese Nachbearbeitung ist. Wenn ich weiß, dass es wiederholt wird, arbeite ich automatisch  
228 ein bisschen genauer. Und wenn diese Nachbearbeitung nur archiviert wird und ich weiß, dass



229 das nie wieder verwendet wird, dann gehe ich nicht ganz so genau vor, und füge nicht jeden  
230 Nebensatz auch noch hinzu.

231 *F: Haben Sie sonst noch Anmerkungen zum Workflow beim Live-Untertiteln?*

232 A: Es ist, finde ich, spannend, weil es sehr unterschiedlich ist von Sendung zu Sendung. Man  
233 weiß eigentlich nie, was einen vorher erwartet. Auch wenn man diverse Sendungen kennt, ist  
234 es doch jedes Mal eine Überraschung, ob die Texte stimmen, wer zu Gast ist, wer moderiert.  
235 Wie rechtzeitig kommt das Material? Also man braucht dennoch eine gewisse Flexibilität, auch  
236 wenn scheinbar immer dasselbe passiert. Und die Erfahrung hilft. Wenn man diverse Modera-  
237 tor\*innen kennt, kann man sich darauf einstellen, ob die Textqualität passt und man schon Sa-  
238 chen kürzen sollte, wenn man weiß, dass die Moderator\*innen schnell sprechen. Aber sonst ist  
239 es immer gut, wenn die Technik funktioniert. Die meisten Zeit verliert man eigentlich, wenn  
240 irgendwelche Programme nicht funktionieren oder ähnliches.

## Abstract – Deutsch

„Respeaking“ bedeutet die simultane Erstellung von (meist intralingualen) Untertiteln mithilfe von Spracherkennung. Da es sich hierbei um einen recht jungen Bereich der Translationswissenschaft handelt, bestehen bis dato kaum wissenschaftliche Untersuchungen hinsichtlich der Arten der live-untertitelten Sendungen und dem dabei entstehenden Workflow.

Um einen Einblick in die Arbeit von Live-Untertitler\*innen zu geben, wurde hierfür eine Aufstellung der Sendungen vorgenommen, die live untertitelt werden, und in einem eigens angelegten Raster festgehalten, welches die Sendungen in vier redaktionelle Programmsparten (Informationssendungen, Sportsendungen, nonfiktionale Unterhaltung, fiktionale Unterhaltung) und die dazugehörigen Sendungsformate gliedert. Außerdem wurde mithilfe eines Beobachtungsrasters zur Selbstreflexion der Workflow einer konkreten Sendung dokumentiert und mit den Erkenntnissen von Expert\*innen-Interviews mit Respeaker\*innen abgeglichen und ergänzt. Zudem wurden problemauslösende Faktoren während des Arbeitsauflaufes festgehalten.

Die Analyse ergab, dass Respeaker\*innen vor allem Informationssendungen wie Magazinsendungen und Nachrichtenprogramme bearbeiten. Sportsendungen und nonfiktionale Unterhaltungssendungen machen dahingegen nur einen geringen Anteil aus. Am häufigsten werden Untertitel im semi-live Modus (v.a. bei Nachrichtensendungen inkl. Wetterbericht) gesendet. Fast ein Drittel aller Sendungen weisen gemischte Untertitel (v.a. bei Magazinsendungen) auf. Nur die wenigsten Sendungen (v.a. im nonfiktionalen Unterhaltungsbereich und beim Sport) werden rein live untertitelt.

Der Workflow in der Live-Untertitelung wurde in Vorbereitung, Live-Sendung und Nachbearbeitung eingeteilt. Die Ergebnisse zeigen, dass je vollständiger und korrekter Material im Voraus zur Verfügung steht, umso mehr erleichtert es die Arbeit der Respeaker\*innen in den drei Phasen. Des Weiteren lässt sich festhalten, dass – neben den Fertigkeiten im simultanen (intralingualen) Dolmetschen – die technische Komponente für das Live-Untertiteln eine wichtige Rolle spielt. Die Respeaker\*innen müssen nämlich (meist gleichzeitig) verschiedene Programme bedienen. Die Untersuchung weist zudem auf, dass neben den bekannten Problemfaktoren beim Simultandolmetschen auch noch technische Faktoren beim Respeaking eine Herausforderung darstellen.

## **Abstract – Englisch**

"Respeaking" describes the simultaneous production of (mostly intralingual) subtitles using speech recognition. As this is a rather young field of translation studies, there is hardly any scientific research on the types of live-subtitled broadcasts and the resulting workflow.

In order to provide an insight into the work of live subtitlers, a list of live subtitled broadcasts was compiled and recorded in a specially created grid. This grid divides the programs into four editorial program categories (information programs, sports programs, non-fictional entertainment, fictional entertainment) and the corresponding program formats. In addition, the workflow of a specific broadcast was documented with the help of an observation grid for self-reflection and compared and brought to completion with the findings of expert interviews with respeakers. Furthermore, problem-triggers during the workflow were recorded.

The analysis showed that respeakers mainly work on information programs such as magazine shows and news programs. Sports programs and non-fictional entertainment programs, on the other hand, account for only a small share. Subtitles are most frequently broadcast in semi-live mode (especially valid for news programs including weather reports). Almost a third of all broadcasts feature mixed subtitles (especially magazine shows). Only very few programs (especially in the non-fictional entertainment sector and in sports) are subtitled completely live.

The workflow in live subtitling was divided into preparation, live broadcast and post-production. The results show that the more complete and correct available material is in advance, the more it facilitates the work of the respeakers in the three phases. Furthermore, it can be stated that – in addition to the skills in simultaneous (intralingual) interpreting – the technical component for live subtitling plays an important role. The respeakers have to operate (mostly simultaneously) different programs. The study also shows that besides to the known problem factors in simultaneous interpreting, technical factors also pose a challenge in respeaking.